UNIVERSITÀ
DI PAVIA

FACULTY OF ENGINEERING

DEPARTMENT OF INDUSTRIAL AND INFORMATION ENGINEERING

MASTER'S DEGREE IN BIOENGINEERING

MASTER'S THESIS

U-NET-BASED DEEP LEARNING SEGMENTATION OF SINGLE-LEAD SMARTWATCH ECGs FOR AUTOMATED PR AND QTc INTERVAL ESTIMATION IN CANCER PATIENTS

SEGMENTAZIONE DI ECG DA SMARTWATCH A SINGOLA DERIVAZIONE BASATA SU DEEP LEARNING CON U-NET PER LA STIMA AUTOMATIZZATA DEGLI INTERVALLI PR E QTc NEI PAZIENTI ONCOLOGICI

Candidate: Miriam Dilda (ID: 538586)

Supervisor: Prof. Pietro Cerveri
Co-supervisor: Prof. Francesco Renna

Academic Year: 2024/2025

Abstract

Cancer therapies are increasingly associated with cardiovascular complications, including atrial fibrillation (AF), which may remain clinically silent while increasing morbidity and mortality. Early detection of conduction abnormalities and repolarization changes is clinically important. Although the 12-lead electrocardiogram (ECG) is the gold standard, wearable devices recording single-lead ECGs offer opportunities for continuous remote monitoring. However, interval measurements from wearable ECGs are not fully validated. We developed and evaluated a transparent, clinically interpretable deep learning (DL) framework for automated estimation of PR and corrected QT (QTc) intervals from single-lead ECGs in oncology patients. Daily recordings over 12 weeks from 40 lung cancer patients were analysed using a U-Net–based DL architecture and compared with the NeuroKit2 algorithm. The method showed good accuracy for PR (mean absolute error (MAE) 14.3 ms vs 24.5 ms) and QTc (MAE 13.6 ms vs 19.1 ms), and smartwatch ECGs agreed closely with paired 12-lead recordings. These findings support automated cardiac interval quantification from wearable ECGs and their potential in scalable remote cardiac monitoring in oncology patients at increased arrhythmic risk.

Keywords: single-lead ECG, deep learning, U-Net, PR interval, QTc interval, wearable ECG, cardio-oncology, atrial fibrillation.

Abstract in lingua italiana

Le terapie oncologiche sono sempre più associate a complicanze cardiovascolari, tra cui la fibrillazione atriale (FA), che può rimanere clinicamente silente pur aumentando morbidità e mortalità. L'identificazione precoce di anomalie della conduzione e alterazioni della ripolarizzazione è quindi importante. Sebbene l'ECG a 12 derivazioni sia il gold standard, i dispositivi indossabili che registrano ECG a singola derivazione offrono opportunità per il monitoraggio remoto continuo, ma le misurazioni degli intervalli non sono completamente validate. Questo studio ha sviluppato e valutato un framework di deep learning (DL) trasparente e interpretabile per stimare automaticamente gli intervalli PR e QT corretto (QTc) da ECG a singola derivazione in pazienti oncologici. Registrazioni giornaliere per 12 settimane di 40 pazienti con carcinoma polmonare sono state analizzate con un'architettura DL basata su U-Net e confrontate con NeuroKit2. Il metodo ha mostrato buona accuratezza per PR (errore medio assoluto (EMA) 14.3 ms vs 24.5 ms) e QTc (EMA 13.6 ms vs 19.1 ms), con buono accordo tra smartwatch e ECG standard.

Parole chiave: ECG a singola derivazione, deep learning, U-Net, intervallo PR, intervallo QTc, ECG indossabile, cardio-oncologia, fibrillazione atriale.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AF Atrial Fibrillation

ECG Electrocardiogram

ML Machine Learning

DL Deep Learning

DNN Deep Neural Network

CNN Convolutional Neural Network

RR Time Interval Between Successive R Waves

QT Time Interval Between Q and T waves

PR Time Interval Between P and R waves

CTIs Cardiac Time Intervals

U-Net U-Net Neural Network

NSR Normal Sinus Rhythm

RF Random Forest

WDs Wearable Devices

SA Sinoatrial

AV Atrioventricular

MAE Mean Absolute Error

1| Introduction

1.1. Context and Motivation

Nowadays cancer is a very prevalent illness, accounting for about a 20 million of new cases in the 2022, according to updated estimates from the International Agency for Research on Cancer (IARC) [13]. In particular, lung cancer is the most frequently diagnosed cancer in 2022, responsible for almost 2.5 million of new cases, or one in eight cancers worldwide (12.4% of all cancers globally) [13]. Oncological patients are treated with various types of therapy, such as chemotherapy and radiotherapy. Both chemotherapy and radiotherapy have been shown to potentially induce cardiac arrhythmias and electrophysiological disturbances, representing a real complication of cancer treatments [14]. In this context, arrhythmias such as atrial fibrillation (AF) are commonly observed during cancer therapies [15], and are associated with worse clinical outcomes [16]. AF is the most common sustained and serious cardiac arrhythmia in humans [17]; indeed, it affects 1-2% of the global population [18]. As in the case for arrhythmias in general, AF consists of an alteration of heart's electrical activity. Detection of AF is a hot topic because it increases the risk of heart failure, dementia, stroke, myocardial infarction, and other-related complications, resulting in a high rate of morbidity and mortality [19], [20]. Moreover, AF often produces no symptoms and may go unnoticed by a patient; for this reason, a reliable method is needed to detect early this type of disorder.

1| Introduction

Continuous monitoring allows early detection and timely intervention, reducing the risk of severe complications, preventing therapy interruptions, and improving overall patient safety and prognosis.

Electrocardiogram (ECG) is considered the gold standard for the AF detection. Typically, an expert cardiologist can identify abnormal heart activities in ECG, which reflects the health status of the heart and helps classify arrhythmias. In controlled environments, such as hospitals or clinical facilities, 12-lead ECG is often used. At the same time, the high technological development has enabled significant progress in wearable devices (WDs), allowing ECG acquisition with only one lead: single-lead ECG. WDs for ECG data acquisition emerged in 1980s, and in recent years, have become more popular. Additionally, they are able to collect single-lead ECG wave comfortably, even when the patients are not in the hospital [21]. Surveys among cardio-oncology specialists have shown that mobile and WDs are considered useful for monitoring cardiac rhythm in active cancer patients, including AF detection. This shows that continuous at-home monitoring is both feasible and recognized in clinical practice [22]. The most important perceived barrier to devices implementation is their high cost.

Thanks to the rapid evolution of technology, it is now possible to implement a continuous home-monitoring approach using single-lead ECG recordings. However, so far, there are still limited established results regarding the quality of cardiac intervals extracted from single-lead ECG, that are relevant for AF research.

In addition, the literature review shows that existing projects focus both on AF detection and AF prediction. These approaches are generally based on end-to-end deep learning (DL) models, often considered as “black box” systems. For this reason, there is still a lack of methods that are more closely related to clinical practice, namely approaches that are more explainable, transparent, and aligned with how clinicians interpret ECG data. In clinical settings, this interpretation is typically based on specific time intervals (PR or corrected QT, QTc), which are commonly adopted as biomarkers for AF.

Moreover, within the field of DL, only a limited number of studies address scenarios with very scarce data. Few approaches combine the use of feature extraction from ECG signals with a DL model, instead of directly applying a DL model to raw data.

These research gaps represent the main motivations that led to the definition and development of this work.

1.2. Objectives

The general aim of this project is to address the research gaps outlined in the previous section, which have not been fully resolved from a scientific perspective yet. To achieve this goal, several aspects related to the analysis and processing of ECG signals have been explored.

The main idea is the development of an innovative DL algorithm for heartbeat segmentation, which allows the extraction of characteristic cardiac events from an ECG signal. These events can then be used to calculate cardiac intervals and other relevant biomarkers for the study of AF in a transparent and explainable manner.

Building on this approach, the specific objectives of this project are as follows:

- Develop an innovative deep learning algorithm for heartbeat segmentation, allowing the extraction of characteristic cardiac events.
- Compute cardiac time intervals (PR and QTc) from the segmented events.
- Combine ECG feature extraction with a DL model to improve robustness when handling low-quality or highly variable signals.
- Ensure reliable ECG signal assessment from WDs, improving accuracy, speed, and robustness compared to traditional methods.
- Evaluate the potential of continuous wearable ECG monitoring for the detection of cardiotoxicity in lung cancer patients.

2| Background

In this chapter, we introduce the heart, focusing on its anatomy and blood circulation in relation to AF. We also discuss the gold-standard technique used to detect cardiac activity: the ECG, and the use of cardiac time interval (CTI) measurements, which are useful for distinguishing between patient groups with or without AF.

2.1. Anatomy of the Heart

The heart is the most crucial organ of the circulatory system. It pumps blood to all parts of the body, providing oxygen and nutrients for their proper function [1]. Especially, the heart works as a double pump [23], ensuring blood circulation through the pulmonary and systemic pathways. Indeed, the heart works as the central organ, inside an intricate network of tubes called blood vessels, which carry blood to and from all parts of the body [1]. Pulmonary circulation is the pathway through which blood travels from the heart to the lungs, where gas exchange takes place: oxygen enters the blood while carbon dioxide is removed. On the other hand, the systemic circulation is the pathway that carries blood from the heart to the rest of the body, delivers oxygen to cells, and removes carbon dioxide, as shown in Figure 2.1.

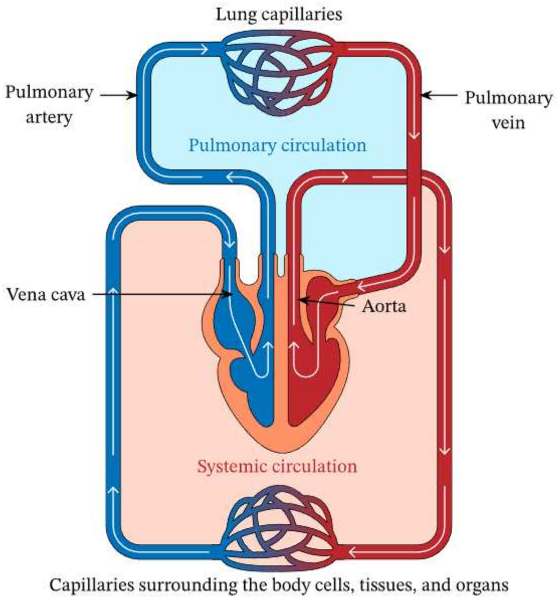

Figure 2.1: A diagram of the pulmonary and systemic circulation [1].

This double circuit ensures continuous oxygenation and an adequate distribution of substances to support cellular functions. Examining the heart's anatomy helps to understand its structure. As shown in Figure 2.2, the heart has four chambers: the upper chambers are called the left and right atria, while the lower chambers are called the left and right ventricles. A muscular wall, called the septum, separates the left and right atria, and the left and right ventricles. This is because the heart is divided in the right part and left part. The right part is involved in the pulmonary pathway, while the left part is involved in the systemic pathway.

The blood flows through the heart via four valves (Figure 2.2) [2], which ensure unidirectional flow between the heart chambers and the vessels: the tricuspid valve regulates blood flow between the right atrium and the right ventricle; the polmonary valve lets blood flow from the right ventricle into the pulmonary arteries; the mitral valve regulates blood flow between the left atrium and the left ventricle; the aortic valve opens the way for oxygen-rich blood to pass from the left ventricle into the aorta, the body's largest artery.

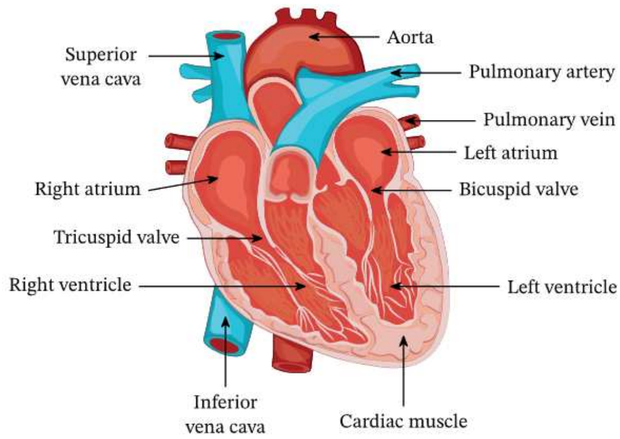

Figure 2.2: A diagram of the human heart and its main blood vessels [1].

2.2. Conduction System

Besides the heart's anatomy, the conduction system is crucial for the correct flow of blood. It consists of electrical impulses from the heart muscle (the myocardium) that trigger the heart's contraction [2]. It is important to note that both the anatomy and conduction system are relevant for AF, because the left atrium undergoes structural and functional modifications at multiple levels; cellular, molecular, and tissue-level. Therefore, AF isn't merely an abnormal rhythm but can also be both a cause and a consequence of anatomical and structural alterations [24]. The intrinsic mechanism of cardiac rhythmicity allows continuous trasmission of electrical impulses that coordinate the contraction of the heart muscle. This property is essential to ensure an effective and continuous heartbeat, fundamental for body's circulatory function. The impulse-conducting system includes the following components:

- sinoatrial (SA) node
- atrioventricular (AV) node
- atrioventricular bundle of His
- right and left branches of the atrioventricular bundle
- subendocardial fibres of Purkinje

Normally, the heart's pacemaker is located in the SA node. From there, the electrical impulse spreads across the atria and reaches the AV node, which slows conduction, causing the physiological delay between atrial and ventricular activation, which appears on the ECG as the PR interval.

From the AV node, the impulse travels through the bundle of His, which enters the upper part of the interventricular septum and then splits into the right and left branches. These branches further divide and form the Purkinje network, which spreads throughout both ventricles, ensuring coordinated contraction of the heart muscle (Figure 2.3).

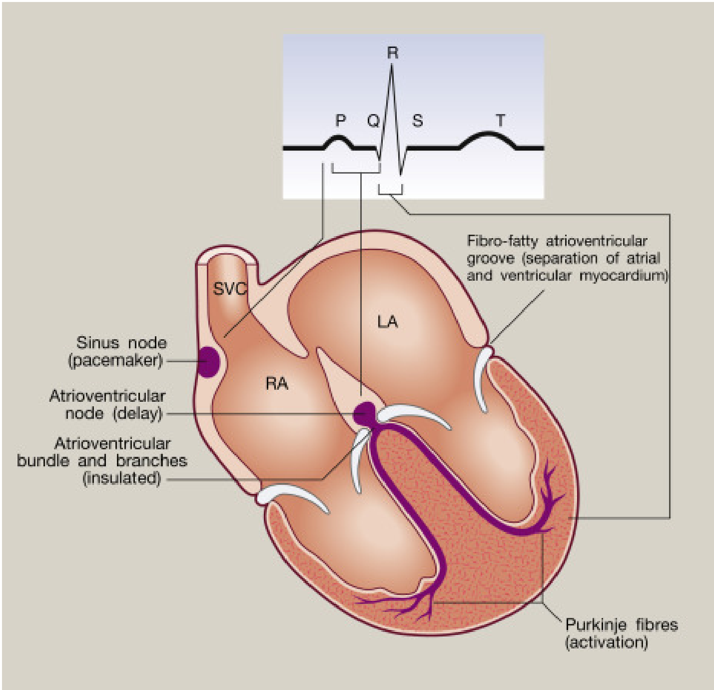

Figure 2.3: Basic structure of the cardiac conduction system and its relation to the chambers and ECG cycle [2].

As mentioned a few lines above, AF significantly impacts the blood circulation. Structural alterations in the left atrium caused by AF interrupt normal electrical conduction and reduce the efficiency of blood flow throughout the body. Mechanical and quantitative studies [25] show that AF can cause tissue hypoperfusion, potentially leading to ischemia, organ dysfunction, and increased vascular risk. To detect AF, which involves both electrical and mechanical disturbances, the ECG remains the standard diagnostic tool.

2.3. Electrocardiogram and AF Detection

The ECG is the gold standard tool used for assessing heart activity. The traditional ECG consists of 12-leads, obtained using 10 electrodes, providing an accurate and robust evaluation. It is usually performed in a controlled environment such as a hospital or clinic. The 12-lead ECG allows a 3D overview of the heart's electrical activity and enables the diagnosis cardiac conditions as infarction [26], [27]. However, with the technology development, it is now possible to record ECG signals using WDs, such as smartwatches, which permit continuous monitoring of cardiac activity at home or during physical exercise. In this case, the recording involves a single-lead with which the spatial information is reduced and the signal is more prone to artifacts noise or insufficient skin preparation [28]. In this study, we analyze electrocardiographic signals acquired from single-lead ECG.

From the ECG signal, it is possible identify five fiducial points that are useful for the analysis. These fiducial points correspond to specific wave peaks: P, Q, R, S and T. Figure 2.5.

The ECG waveform reflects a series of coordinated electrical events within the heart's conduction system, representing the combined electrical signals.

These signals originate from the propagation of action potentials across cardiac cells, leading to changes in their membrane potential [29].

P wave represents atrial depolarization and the onset of atrial contraction. The cardiac cycle begins with the firing of the SA node, which is not visible on a typical ECG. Its depolarization spreads rapidly through both atria, generating the P wave. In contrast, when AF is present, the organized electrical activity originating from the SA node is replaced by chaotic and disorganized atrial impulses. As a consequence, the P wave is absent and is replaced by fibrillatory waves (F-waves - see Figure 2.4), which are continuous, irregular, low-amplitude atrial waveforms on the baseline [30], [31].

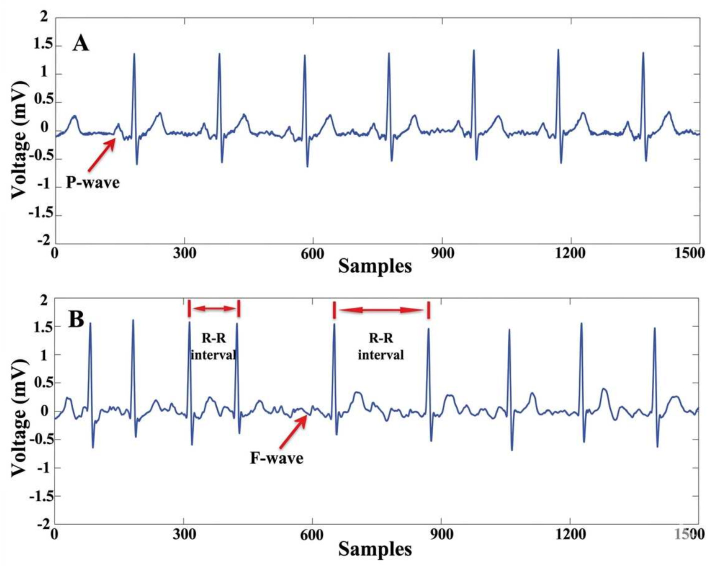

Figure 2.4: Comparison of ECG features between (A) a normal subject and (B) a patient with AF, showing AF characteristics in (B) [3].

QRS complex represents ventricular depolarization and the onset of ventricular contraction. Following the P wave, the ventricles depolarize, producing the QRS complex. The first negative deflection is the Q wave, the large positive deflection is the R wave, and the subsequent negative deflection is the S wave. When the QRS complex ends, the ventricles are fully depolarized and contracting, while the atria have finished contracting and are repolarizing.
T wave represents ventricular repolarization. As the ventricles repolarize after contraction, the ECG signal returns to baseline, forming the T wave. It is the last detectable potential in the cardiac cycle, followed by the P wave of the next cycle.

2.3.1. CTIs extraction for AF

From the ECG and its fiducial points, it is possible to extract multiple cardiac intervals. However, our focus will be on two specific intervals: PR and QT (or QTc), as they are particularly relevant in the study of AF.
The **PR interval** represents the time between the onset of atrial depolarization (start of the P wave) and the onset of ventricular depolarization (beginning of the QRS complex). Indeed, this interval can also be referred to as the PQ interval [32]. Epidemiological studies have shown that a prolonged PR interval is significantly associated with an increased risk of AF. Analyses indicate that a prolonged PR is linked to a higher likelihood

of developing AF, confirming its predictive role in identifying individuals at risk [33]. In a healthy population, the PR interval is typically less than 200–212 ms, according to Olbertz et al., [34].

The **QT interval**, particularly the **QTc** (QT-corrected for RR interval), reflects the total duration of ventricular depolarization and repolarization. Studies have shown that a prolonged QT interval is associated with an increased risk of incident AF. Notably, each 10 ms increase in QTc corresponds to a proportional rise in risk, indicating that even relatively small changes in this interval can have clinical significance [35]. In a healthy population, the QTc, corrected using the Fridericia formula

$$QT_c = \frac{QT}{\sqrt[3]{RR}} \tag{2.1}$$

is generally below 450 ms [34].

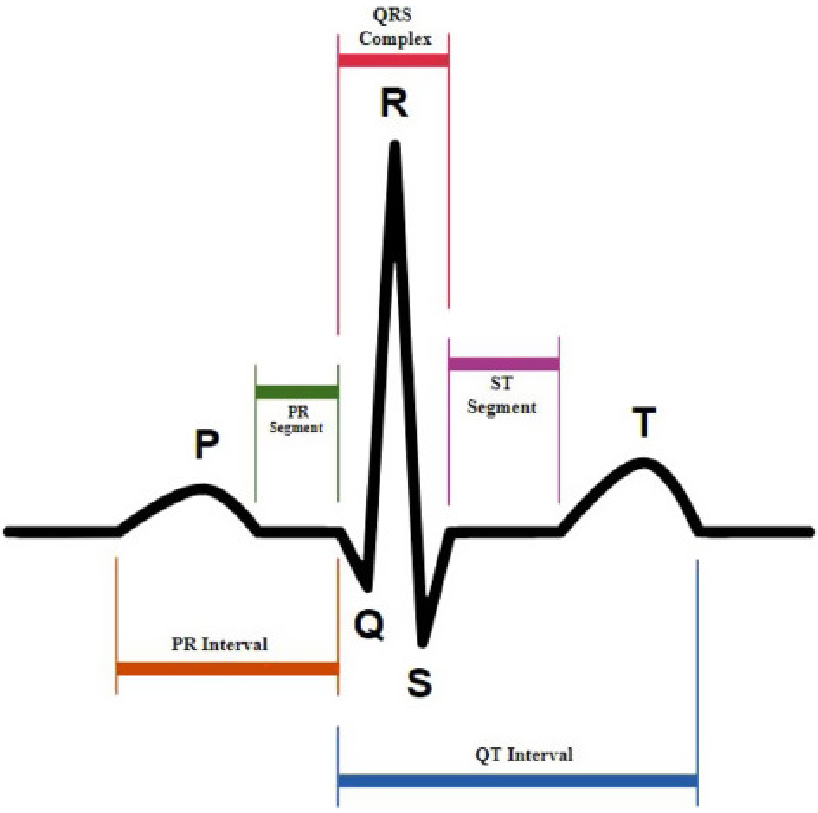

Figure 2.5: P, QRS, T waves, PR interval, and QT interval on ECG signals [4].

CTIs, derived from ECG fiducial points, provide valuable measures for studying cardiac electrical activity and for identifying potential predictors of AF, particularly by focusing on PR and QTc intervals.

3 | Literature Review

3.1. Research Aim

Once the problem is fully understood, the next step of this thesis is to investigate how researchers in recent years have addressed the same issue; namely, to identify which biomarkers can be obtained from ECG signals, particularly single-lead ECG recordings, that are related to AF, and to understand the methods and algorithms used to obtain them. Among the various biomarkers, the QT interval and the P wave represent two of the most widely used and clinically relevant features to be analyzed.
The aim is to highlight methodologies that have been implemented and compare the performance of previous studies. The research question that defines the scope of this literature review section has therefore been carefully formulated as follows: "Which methodologies have been developed to monitor the condition of a patient in real time using a single-lead ECG?"

3.2. Review of Atrial Fibrillation's Detection

In recent years, many new approaches have been developed for the detection of AF, with the aim of improving patient outcomes. Thanks to the rapid development of digital technologies, in particular artificial intelligence, machine learning and deep learning, these methods have become increasingly effective for medical signal analysis. Consequently, numerous studies have contributed to the expansion of both scientific and clinical knowledge in this field. In this part of literature review, the focus is on understanding how previous studies have achieved AF detection and on comparing and analyzing the strategies used to build reliable models for classifying AF.

The main aspects that differentiate these approaches, and therefore allow meaningful comparison, include: the dataset used for the investigation, the computational approach adopted (ML, DL or statistical methods) and the features extracted from the ECG data for AF detection.

Based on the overall literature analysis, this section summarizes the most relevant studies, highlighting the aspects that are most pertinent to this work. It should be noted that the majority of the identified studies focus on AF detection rather than prediction. Several authors have investigated AF using single-lead ECG signals, primarily through ML and DL approaches. However, statistical analysis methods have also been explored, albeit to a lesser extent. Among these methodologies, DL based approaches appear to be dominant. CNNs represent the most commonly employed architecture within the DL domain. CNNs can be applied to data with grid-like structure (2D images) or even sequences (1D time domain signals). Their sparse connectivity and parameter sharing properties allow them to automatically learn discriminative features from these kinds of data. In addition, some approaches transform ECG signals into 2D time-frequency representations, which can be processed as images by the CNNs. Within this framework, two studies are particularly noteworthy: Lai et al., 2019 [3] and Alam R et al., 2023 [5].

3.2.1. Lai et al., 2019

Lai et al. [3] proposed a model based on two parallel CNNs designed to distinguish between AF and non-AF signals. The outputs of the two networks are subsequently combined to improve overall performance. Specifically, the model uses two input channels: the RR interval series for the first channel and the F-wave frequency spectrum for the second.

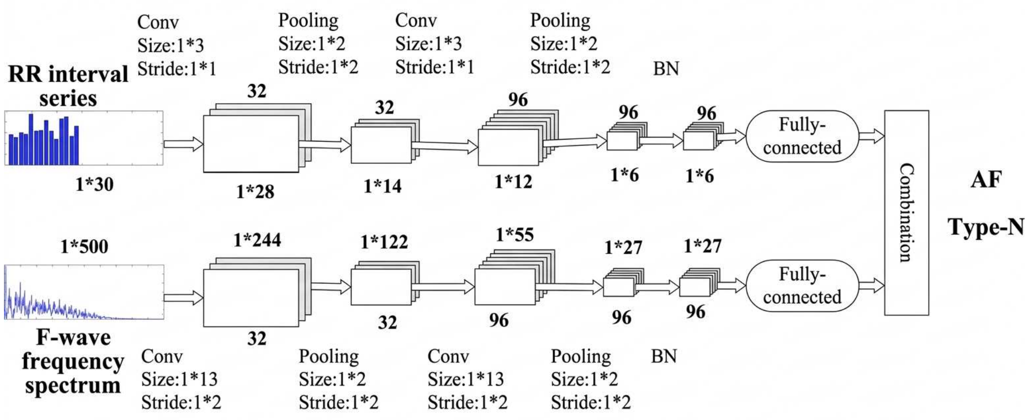

Figure 3.1: Illustration of the proposed lightweight CNN by using representative cardiac rhythm features, such as the RR intervals and F-wave frequency spectrum, as input rather than raw ECG signals [3].

In contrast to approaches based on deeper architectures, this method employs a lightweight CNN that takes RR interval series and F-wave frequency spectra as inputs instead of raw ECG signals, without relying on explicit electrophysiological assumptions. Although the network is relatively shallow, it achieves a final accuracy of 97.5%.

3.2.2. Alam R et al., 2023

Alam R et al., 2023 [5] proposed an alternative CNN-based application. Unlike traditional classification tasks, their objective is the estimation of QT intervals and heart rate using a single-lead ECG derived from a 12-lead ECG dataset, given that QT prolongation can often lead to fatal arrhythmias. The proposed model, named QTNet, is based on a residual neural network architecture, as illustrated below.

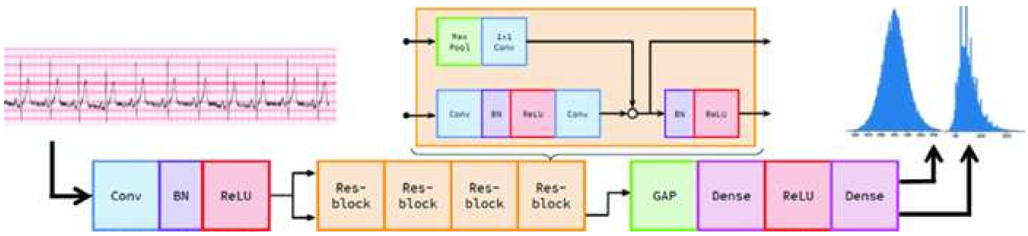

Figure 3.2: QTNet architecture, input: Lead-I ECG (from 12-lead ECG) sampled at 250 Hz, output: QT interval and heart rate [5].

For QT interval regression, the model was evaluated on four external datasets, including publicly available databases. The results demonstrated a strong correlation between the predicted values and the corresponding 12-lead reference labels across all datasets.

The MAE for QT interval estimation ranged between 9 ms and 15.8 ms, while Pearson correlation coefficients varied between 0.899 and 0.914. In comparison, QT interval estimation using a standard automated ECG analysis tool (NeuroKit2 [36]) yielded MAEs between 22.29 ms and 90.79 ms, with Pearson correlation coefficients ranging from 0.345 to 0.620.

3.2.3. A. Luca et al., 2016

Beyond DL-based approaches, it is also relevant to consider statistical methods for AF detection. A. Luca et al., 2016 [6] proposed a statistical approach to classify Normal Sinus Rhythm (NSR) and AF from single-lead ECG signals by examining differences in specific ECG features between subjects.

The study extracted several features based on P-waves and RR intervals. Among these, the most discriminative features for identifying AF susceptibility were P-wave duration, the mean and standard deviation of the beat-to-beat Euclidean distance between consecutive P-waves, and the sample entropy of RR intervals. Notably, the results show that their combined use enables an efficient separation of the two populations and can therefore serve as an effective tool for identifying patients at risk of developing AF.

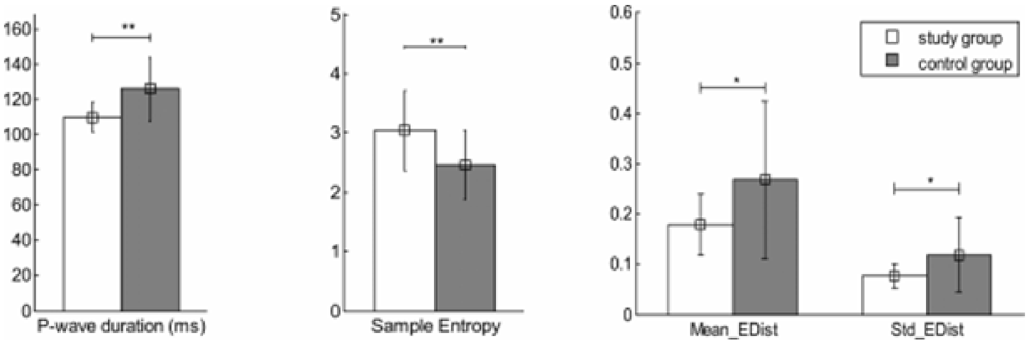

Figure 3.3: ECG characteristics features extracted from the lead II. The mean (square) of each group is shown. The vertical lines delimit the mean ± the standard deviation [6].

The statistical comparison was performed by using the Mann-Whitney U test for unpaired data. Patients with a history of AF exhibited significantly longer P-wave duration and higher mean and standard deviation of the beat-to-beat Euclidean distance compared to the control group. On the other hand, HRV analysis showed that patients with a history of AF had lower sample entropy than patients without previous AF history, meaning that RR intervals exhibited lower complexity in subjects with AF.

3.2.4. Discussion on Research

Although a substantial body of literature addresses AF detection and prediction, relatively few approaches focus on developing models that are transparent and explainable, and therefore closer to clinical practice. The studies highlighted in this review provide several relevant insights.

Lai et al. [3] suggest the use of extracted ECG features as inputs to a CNN. However, this approach requires transforming the signal domain (e.g., from 1D to 2D), which introduces additional preprocessing complexity. From a technical perspective, this study supports the use of CNN architectures with feature-based inputs, although the approach remains largely "black-box" in nature.

Alam R et al., 2023 [5] is relevant from both technical and clinical perspectives. The study addresses QT interval regression, a clinically meaningful parameter in AF research, and demonstrates the feasibility of adapting CNN architectures to operate directly on one domain. A limitation is that the single-lead signal is extracted from a 12-lead ECG, implying higher signal quality compared to a truly single-lead acquisition. Furthermore, Alam R et al.'s regression framework underscores the widespread use of NeuroKit2 as a baseline benchmark for ECG interval estimation, because NeuroKit2 relies on classical signal delineation to determine fiducial points, while deep learning models such as QTNet use an end-to-end regression approach.

Finally, A. Luca et al., 2016 [6] despite not employing DL methods, provide clinically relevant insights by emphasizing the importance of P-wave duration in identifying AF history. This finding supports the consideration of clinically interpretable time intervals, such as the PR interval, as potential biomarkers to be extracted from ECG signals in the present study.

Overall, the reviewed studies highlight existing gaps that this work aims to address: the need to operate effectively with scarse data, to develop more transparent and clinically interpretable approaches, particularly through the analysis of clinically relevant CTIs, and to enable a home-monitoring framework based on single-lead ECG signals.

For a detailed structural analysis of the reviewed literature, see Appendix 8.1.

4 | Materials and Methods

This chapter first describes the dataset and then explains the method used in this study to obtain the final algorithm.

4.1. Dataset

The dataset was collected in collaboration with the Cardiac Department of the Unidade Local de Saúde Gaia e Espinho (ULSGE). It includes 40 lung cancer patients, with 30% females and 70% males, aged between 42 and 80 years.

Variable	**Value**
Female	30%
Male	70%
Age range	42–80 years

Table 4.1: Demographic characteristics of the study population.

All patients underwent oncological treatments such as chemotherapy or radiotherapy. Each patient was provided with a Samsung Galaxy Watch5 for a period of three months, with the aim of recording single-lead ECG signals daily, sometimes more than once per day. These recordings are used to analyze the cardiac activity over time and to evaluate whether pharmacological oncological treatments could induce changes in it, with particular attention to the development of AF. Only two patients (the 3rd

and the 40th) developed AF during the study, which makes the dataset strongly unbalanced.
The dataset contains 6273 ECG single-lead recordings, sampled at 500 Hz, each lasting approximately 30 seconds.
The study was conducted in accordance with the approved ethical protocol (code 83-2024-1). Patients were included if:

- were aged 40 years or older.
- had no history of AF or atrial flutter.
- had signed informed consent.
- were candidates for chemotherapy that included at least one of the following agents: Crizotinib, Ceritinib, Osimertinib, Cisplatin, Carboplatin, Cyclophosphamide, Paclitaxel, Nivolumab, Pembrolizumab, Ipilimumab, Gemcitabine or Pemetrexed.

The 40 patients included in the study presented various histological types and stages of lung cancer. Some patients presented synchronous lung tumors (two or more tumors in the same or contralateral lung). This heterogeneity highlights the diverse clinical profile of the study population, providing a realistic context for evaluating cardiac activity in patients undergoing oncological treatments.

Moreover, clinical information was collected for each patient, including quality of life data, echocardiographic parameters, hemoglobin levels, and other clinical measurements. Importantly, 12-lead ECG recordings were also acquired in the hospital using Mortara 350 or Mortara 250c machines. These recordings were performed on the first day and on the last day (after 90 days) of the study and are used in this study as a reference standard to validate the cardiac parameters extracted from the Samsung smartwatch signals.

4.1.1. Data Quality

All single-lead ECG signals are stored in a `.pkl` file, a Python-specific proprietary format [37], with the following columns: *week* (week of acquisition - 4: first month, 8: second month, 12: third month), *patient name* (patient ID), *file* (path of the file containing the ECG signal), and *ECG Signal* (single-lead ECG amplitude values). After organizing the data into a single file, signal quality assessment was performed. Initially, the first 5 seconds of each signal were discarded in order to remove acquisition artifacts occurring at the beginning of the recording. Then, the quality of the remaining signal is evaluated. Among the available signal quality indices, the pSQI index is selected. The pSQI is defined as the ratio between the power spectrum $P(f)$ of the ECG signal integrated in the 5–15 Hz band and that integrated in the 5–40 Hz band:

$$pSQI = \frac{\int_5^{15} P(f)df}{\int_5^{40} P(f)df} \tag{4.1}$$

According to the pSQI value, each ECG signal can be classified as follows:

$$ECG \begin{cases} optimal & \text{if } \ pSQI \in [0.5, 0.8] \\ suspicious & \text{if } \ pSQI \in [0.4, 0.5] \\ unqualified & \text{if } \ pSQI < 0.4 \ \text{ or } \ pSQI > 0.8 \end{cases} \tag{4.2}$$

In this case, to ensure the highest signal quality, the recording is divided into consecutive windows of 15 seconds. The pSQI is evaluated for each window, and the window with the best signal quality is selected among the available candidates to improve the reliability of subsequent processing steps.

pSQI Class (best window)	**Count**
Optimal	1387
Suspicious	1635
Unqualified	3251

Table 4.2: Distribution of pSQI classes in the best window

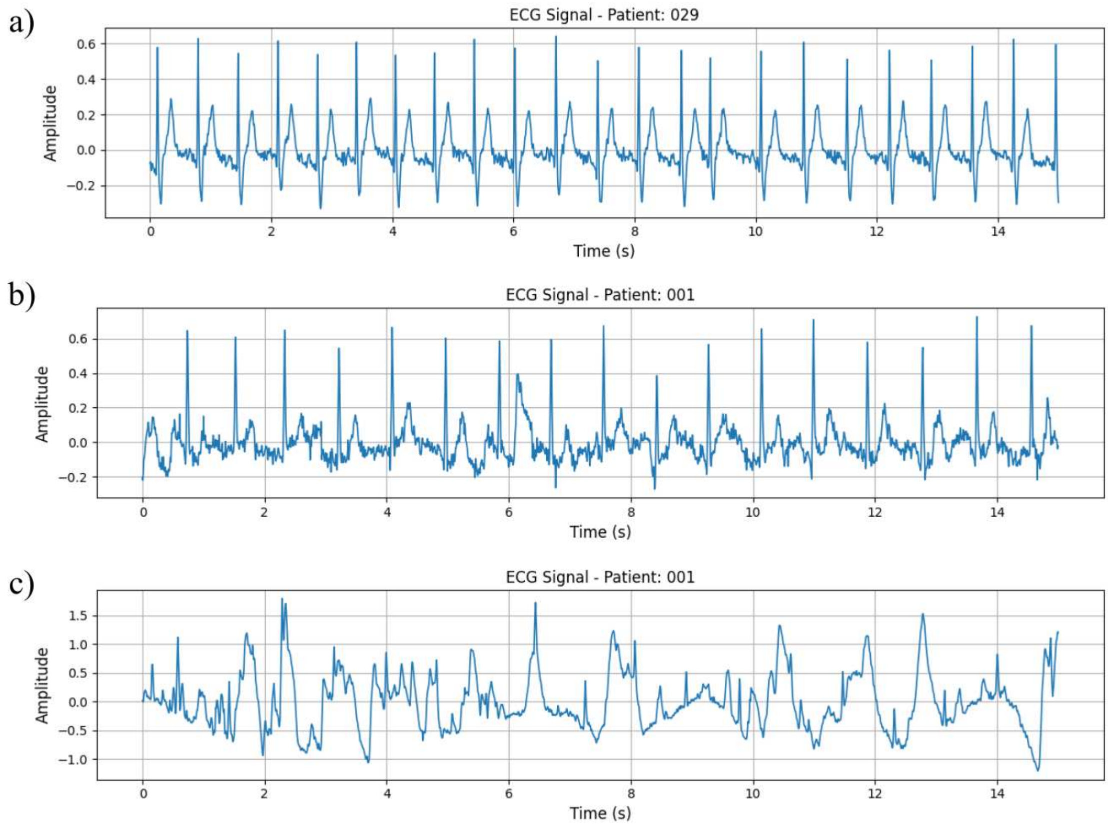

Figure 4.1: Plot of three single-lead ECG signals after quality assessment: a) optimal, b) suspicious, c) unqualified.

Finally, the ECG signal considered for the study and further preprocessed in the next section has a duration of 15 seconds.

4.2. Preprocessing

Before applying any algorithm to the dataset, a preprocessing step is required. The smartwatches used to collect the ECG single-lead signals have private documentation, and details on how they process the signals (for example, any internal filtering) are not publicly available. For this reason, we applied the following preprocessing steps to make sure that the input is consistent and reliable for the next stages. Below, we provide a detailed description of each preprocessing step. First, the dataset organized with the following columns week, patient name, file, ECG signal is standardized using:

- Z-Score Standardization:

$$x_i[n] \longrightarrow \hat{x}_i[n] = \frac{x_i[n] - \mu_{x,i}}{\sigma_{x,i}}$$

This operation ensures $\text{mean}(\hat{x}_i[n]) = 0$ and $\text{std}(\hat{x}_i[n]) = 1$ for each signal. The signals $\hat{x}_i[n]$ were then processed with filtering operations:

1. Notch Filter at 50 Hz: an IIR notch filter with quality factor $Q = 30$ was applied to remove power-line interference at 50 Hz. Then, zero-phase filtering was performed using a forward-backward method:

$$y_i[n] = \text{filtfilt}\big(\{b_k\}, \{a_k\}, \hat{x}_i[n]\big).$$

2. Bandpass Butterworth (0.5–100 Hz): the informative content of an ECG signal is mainly between 0.5 Hz and 100 Hz. Therefore, to study the signal in the relevant range, a Butterworth bandpass filter was applied. This filter keeps only the frequencies within the selected

band, removing all components below 0.5 Hz and above 100 Hz. The signals were filtered using a 6th-order Butterworth filter with $f_{\text{low}} = 0.5$ Hz and $f_{\text{high}} = 100$ Hz.

3. Median Removal: finally, for each signal the sample median is subtracted to remove any DC offset:

$$z_i[n] = y_i[n] - \underset{k=0}{\overset{N_i-1}{\text{median}}}\{y_i[k]\}$$

After filtering, the preprocessed signals were used as input for the NeuroKit2 benchmark and for the developed method, which are explained in detail in the following sections.

4.3. Benchmark Neurokit2

This section focuses on NeuroKit2 [36, 38], a user-friendly Python package that provides easy access to advanced biosignal processing routines.
In this work, NeuroKit2 is used as a reference benchmark to detect temporal events in single-lead ECG signals, from which beat-by-beat cardiac intervals are extracted. After applying the `ecg_process` function, which allows the detection of R-peaks for each heartbeat, the `ecg_delineate` function is used to extract P, Q, S, and T events from the ECG signal.

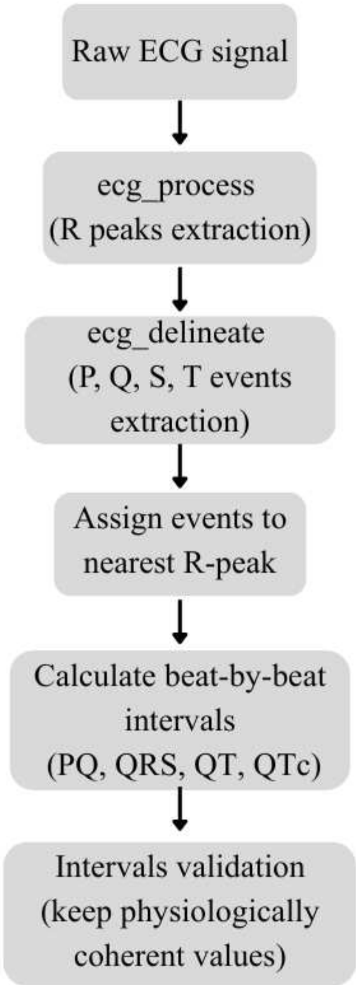

Figure 4.2: Workflow of the NeuroKit2 ECG processing pipeline: R-peak detection, delineation of P/Q/S/T events, beat assignment, cardiac intervals calculation, and selection of physiologically valid intervals.

The extracted events are considered as temporal points, in particular: `P_Onsets` (onset of the P wave), `Q_Peaks` (center of the Q wave), `S_Peaks` (center of the S wave) and `T_Offsets` (end of the T wave).

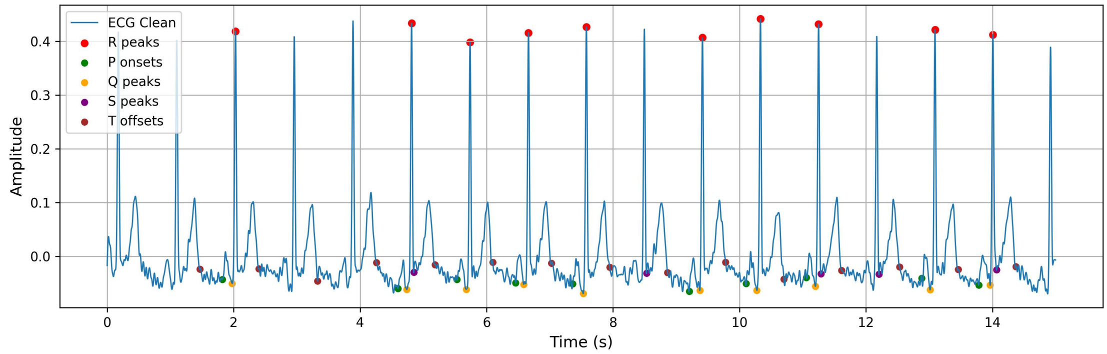

Figure 4.3: Example of NeuroKit2 processing applied to a single-lead ECG signal.

To correctly assign events to each heartbeat, the events closest to the corresponding R-peaks are selected. Cardiac intervals are then calculated as the time differences between events and are considered physiologically valid only if they fall within coherent ranges.

4.4. U-Net-based Beat-by-Beat Model

In this research, the focus is on developing a robust algorithm for the extraction of CTIs from ECG single-lead signals, analyzing the obtained results, and performing a comparative analysis with the benchmark Neurokit2.

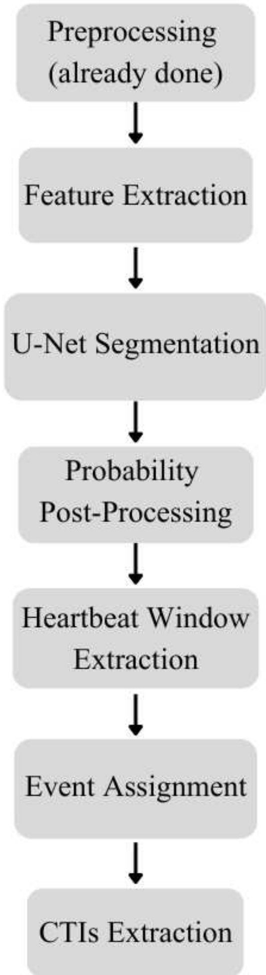

Figure 4.4: Overview of the segmentation pipeline.

The algorithm employs a U-Net model (detailed in Section 4.4.2) to perform beat-by-beat signal segmentation, followed by a post-processing strategy applied to the output probability maps. The model outputs four class probability maps (four channels), and the ECG label sets are defined as follows:

$$\{0 : \text{baseline},\ 1 : \text{P-wave},\ 2 : \text{QRS},\ 3 : \text{T-wave}\}$$

The previous figure illustrates the adopted workflow (see Fig. 4.4).

4.4.1. Features Extraction

The ECG data underwent a preprocessing step, described in Section 4.2, followed by feature extraction before being used as input to the U-Net architecture. This procedure reduces noise and enhances relevant signal characteristics.

The ECG signals are transformed into a set of four time-domain features (or "envelopes"), as follows, where $x(t)$ denotes the input ECG signal and $\epsilon > 0$ is a small constant introduced to avoid numerical instability:

1. Hilbert Envelope: captures amplitude variations by computing the analytic signal using the Hilbert Transform.

$$E_H(t) = |x(t) + j\,\mathcal{H}\{x(t)\}| \tag{4.3}$$

2. Shannon Energy Envelope: highlights high-energy segments by applying a logarithmic energy-based transformation.

$$E_S(t) = \left| -x^2(t)\ \log\left(x^2(t) + \epsilon\right) \right| \tag{4.4}$$

3. Homomorphic Envelope: extracts the signal envelope by computing the logarithm of the absolute signal values, applying a median

filter (window size: 21 samples), and exponentiating the result.

$$E_{Ho}(t) = \exp\left(\text{medfilt}\left(\log\left(|x(t)| + \epsilon\right), 21\right)\right) \tag{4.5}$$

4. Smoothed Envelope: applies a smoothing operation using a 21-sample Hamming window $w(t)$ to reduce high-frequency fluctuations.

$$E_{Sm}(t) = (|x(t)| * w(t)) * w(t) \tag{4.6}$$

All four feature sequences are then downsampled to 50 Hz to reduce the computational complexity of the model. They are stacked to form a $4 \times L^{'}$ matrix and split into patches of 64 samples with an 8-sample stride.

Finally, the ECG feature patches are used as input to the U-Net architecture for signal segmentation.

4.4.2. U-Net Architecture

The segmentation approach employs a U-Net model originally developed by Ronneberger et al. [39] and subsequently customized by the Multiscope Investigation Group at INESC TEC (Porto, Portugal) to perform signal segmentation, with a particular focus on heartbeat segmentation. The model was conceptualized and trained by [40], described in Cardiology Conference 2025, and its architecture was based on [7].
For training, the model takes specific input features and is guided by specific

label sets to achieve accurate segmentation. It was trained on the LUDB dataset [41], using ECG patches and labels {baseline, P-wave, QRS, T-wave}. One reason this dataset was chosen is that it contains signals coming from different leads, so it provides sufficient variability to help the model generalize well, even for non-standard lead morphologies. The LUDB contains recordings from 200 subjects, each contributing a single 10-second, 12-lead ECG (leads I, II, III, aVR, aVL, aVF, V1–V6) sampled at 500 Hz.

The architecture adapted from [7] is structured as follows:

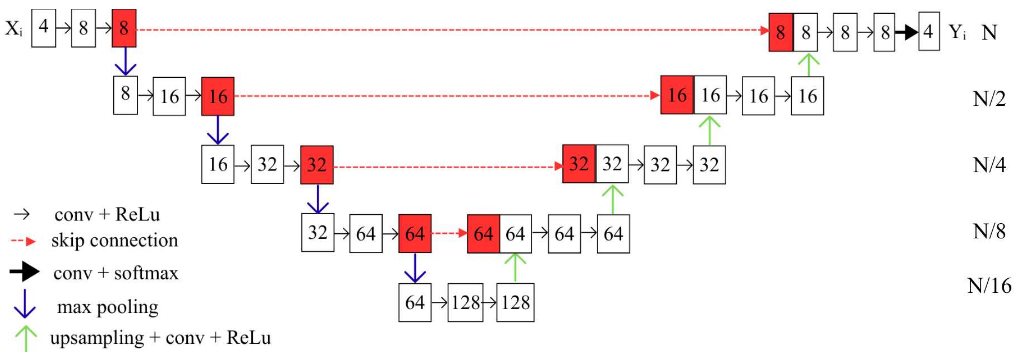

Figure 4.5: U-Net architecture implemented for beat-by-beat segmentation. The numbers inside the boxes represent the number of feature maps; values on the right indicate the temporal length (in samples) for each layer [7].

Input: A 4×64 patch of feature envelopes.

Encoder: Four downsampling blocks, each composed of two 1D convolutional layers (kernel size 3), followed by batch normalization, ReLU activation, and a 2× max-pooling layer. The number of feature maps doubles at each downsampling step ($16 \rightarrow 32 \rightarrow 64 \rightarrow 128$).

Bottleneck: Two 1D convolutional layers with 256 feature maps, batch normalization, and ReLU activation.

Decoder: Four upsampling blocks, each including a transposed convolution (stride 2), concatenation with the corresponding encoder feature map (skip connection), followed by two 1D convolutional layers (kernel size 3), batch normalization, and ReLU activation. The number of feature maps is halved at each upsampling step ($256 \rightarrow 128 \rightarrow 64 \rightarrow 32 \rightarrow 16$).

Output: A final 1D convolutional layer (kernel size 1) generating four channels, followed by a softmax function to obtain class probabilities for each sample in the 64-sample patch.

This design is particularly effective for capturing multi-scale temporal features while maintaining precise localization through skip connections.

4.4.3. Postprocessing U-Net

Once the U-Net neural network has completed its processing, the outputs are probability values. Therefore, additional steps are needed to convert these probabilities into cardiac intervals of interest.

Step 1: Probability Map and Labeling

- The U-Net output is a probability map, where each ECG sample has four probability values:

$$p(t) = \left[p_0(t),\ p_1(t),\ p_2(t),\ p_3(t)\right]$$

- To assign a class to each sample, an argmax operation is applied:

$$\hat{y}(t) = \arg \max_{c \in \{0,1,2,3\}} p_c(t)$$

- This operation generates a sequence of integer labels describing the ECG signal, where each sample is assigned the class with the highest probability.

Step 2: Event Extraction

- Each signal is represented as a list of tuples containing the event type and its corresponding time instant:

$$(\mathrm{P}, t),\ (\mathrm{baseline}, t),\ (\mathrm{QRS}, t),\ (\mathrm{T}, t)$$

- Since multiple consecutive samples may belong to the same event, each event can be identified by its start and end times, enabling the extraction of cardiac intervals of interest.

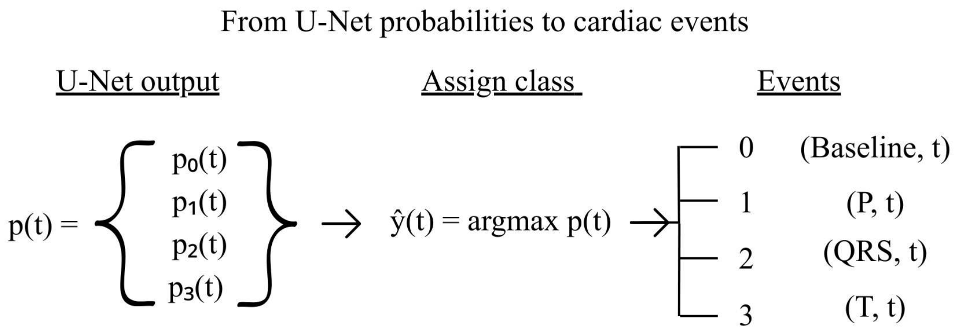

Figure 4.6: Workflow from U-Net output probabilities to discrete cardiac events. Each sample is mapped to an event (Baseline, P, QRS, T) with its corresponding time t.

Step 3: R-Peaks Integration

- After events segmentation, R-peaks are extracted from the signal and added to the event list.
- R-peaks extraction is perfomed on the preprocessed ECG signal using BioSPPy (biosppy.signals.ecg), a module that provides methods to process single-lead ECG signals. In particular, the implemented function is:

  ```
  biosppy.signals.ecg.ASI_segmenter
  ```

  an ECG R-peak segmentation algorithm [8].
- BioSPPy was selected instead of NeuroKit2 because, according to

[8], it provides more reliable R-peak detection without missing a significant number of R-peaks.

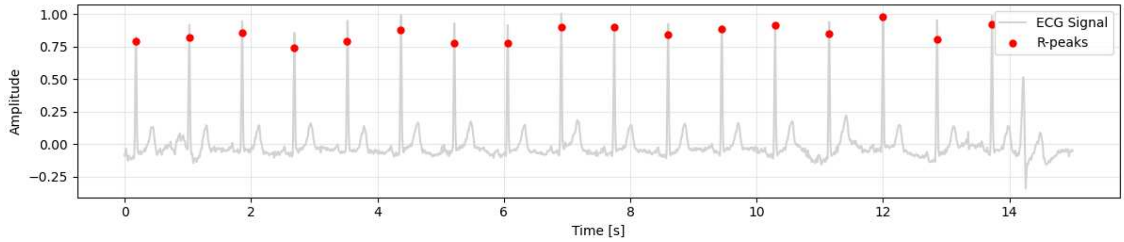

Figure 4.7: R-peaks detection using BioSPPy [8].

Step 4: Heartbeat Window Detection and Event Assignment

The goal of this step is to identify discrete heartbeat windows based on the detected R-peaks. The procedure is as follows:

1. Compute median RR interval:

$$\Delta R_i = R_i - R_{i-1}, \quad i = 2, \ldots, M$$

$$\text{medianRR} = \text{median}\{\Delta R_i\}_{i=2}^{M}$$

   - If only one R-peak is found, medianRR is set to 0.8 s by default.

2. Define asymmetric heartbeat windows:

$$W_{i,\text{start}} = R_i - 0.4 \cdot \text{medianRR}, \quad W_{i,\text{end}} = R_i + 0.6 \cdot \text{medianRR}$$

   - The asymmetry is chosen because the P wave is closer to the R-peak than the T wave, and a symmetric window would not correctly capture the entire heartbeat.

3. Assign events to each heartbeat window:

$$W_i = (W_{i,\text{start}}, W_{i,\text{end}}), \quad W_{i,\text{start}} \leq Time_{event} \leq W_{i,\text{end}}$$

Step 5: Filtering Valid Beats

Only valid beats are retained according to the following criteria:

- Only one R-peak is present within each heartbeat window.
- At least two QRS events must be detected.
- The R-peak must lie between the QRS events segmented by the U-Net.

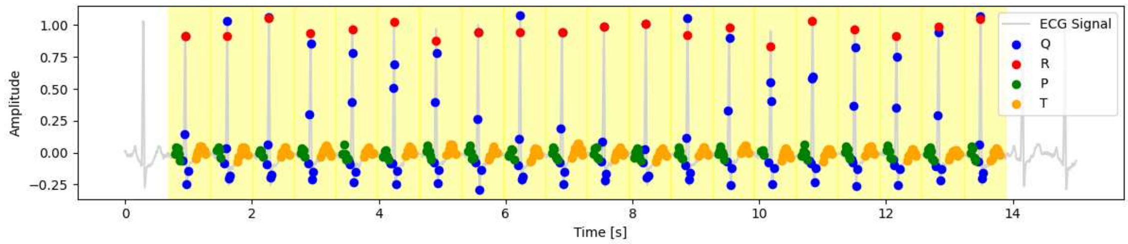

Figure 4.8: Example of filtered beats. The yellow background highlights valid beats and their corresponding events. The Q event represents the QRS label predicted by the U-Net.

Step 6: Cardiac Interval Calculation

For each valid beat, the following intervals are computed:

- PR interval (or PQ interval): the temporal difference between the beginning of the P wave, $\min(P)$, and the beginning of the Q wave, $\min(Q)$:
$$PR = \min(Q) - \min(P)$$

- QT interval: the temporal difference between the beginning of the Q wave, $\min(Q)$, and the end of T wave, $\max(T)$:
$$QT = \max(T) - \min(Q)$$

- QTc interval: QT corrected for the current RR interval using Fridericia formula (see Equation 2.1). The current RR is computed as the time difference between the current R-peak and the R-peak of the previous heartbeat.

Step 7: Mean of Intervals

- For each signal, the final cardiac interval values are computed as the mean across all detected heartbeats. Initially, the median was used; however, after comparing the results obtained with the mean, it was observed that outlier values did not have a substantial influence on the estimates. For this reason, the mean was ultimately adopted.

4.5. Statistical Analysis of CTIs Estimation

In this study, to evaluate the accuracy of the CTIs estimated from the single-lead ECG signal, we computed the MAE and the Pearson correlation coefficient (r). These metrics were calculated for the CTIs obtained from the single-lead ECG using the U-Net neural network and for those obtained with NeuroKit (our benchmark), both compared with the reference CTIs derived from the 12-lead ECG data.

This comparison allows us to assess whether our method provides improved accuracy relative to NeuroKit.

4.5.1. Evaluation Methodology

To evaluate the performance of the CTIs, the MAE and the Pearson correlation coefficient (r) were computed on a validation set consisting of a subset of the available database, including only four signals.
MAE:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - y_{r,i}\right| \tag{4.7}$$

Legend:

- x_i = interval estimated from the single-lead ECG signal
- $y_{r,i}$ = reference interval from the 12-lead ECG signal (ground-truth)
- n = total number of signals

The MAE measures the average absolute difference between the estimated and reference intervals, providing a direct measure of estimation accuracy.

Pearson Correlation Coefficient (r):

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_{r,i} - \bar{y}_r)}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_{r,i} - \bar{y}_r)^2}} \tag{4.8}$$

Legend:

- x_i = CTI estimated from the single-lead ECG signal
- $y_{r,i}$ = reference CTI from the 12-lead ECG signal (ground-truth)
- $\bar{x}$ = mean of the estimated CTIs
- $\bar{y}_r$ = mean of the reference CTIs
- n = total number of signals

The Pearson correlation coefficient quantifies the strength and direction of the linear relationship between estimated and reference intervals. Values close to $+1$ indicate a strong positive linear correlation.

The limited number of validation signals is due to the availability of appropriate reference data. The evaluation was conducted using two different approaches:

1. Using the reference CTI values automatically computed by the Mortara device at the hospital and stored in the hospital 12-lead ECG

database.

2. Using the reference CTI values obtained from the output of our model when the corresponding 12-lead ECG signals were used as input.

In the first approach, a 12-lead ECG was recorded for each patient at two time points: on the first day of the study and after 90 days.
In the second approach, however, only four patients had a 12-lead ECG recorded immediately after the single-lead acquisition performed with the smartwatch. These 12-lead signals were acquired using the Mortara device and are temporally very close to the corresponding single-lead recordings. For this reason, they are particularly suitable for validation purposes. However, this also limits the analysis, since only four such paired recordings (single-lead and corresponding 12-lead ECG) are available.

Because the second approach involves only four paired signals, the first approach was aligned for evaluation using exactly the same four single-lead signals. Specifically, for each of these signals, the reference CTI values were selected from the hospital 12-lead ECG database at the time points closest to the single-lead acquisition. Depending on availability, these correspond either to recordings from the first day of the study or from the 90-day follow-up, with a maximum delay of a few days between the single-lead and the selected 12-lead ECG.

Overall, to ensure a fair comparison, the MAE and Pearson correlation were computed considering only these four corresponding pairs of signals.

The use of both approaches is necessary because, in the second approach, there is a potential risk of comparing results that may share model-specific errors, since both estimates are derived from the same model. However, this approach provides more reliable comparisons in terms of physiological consistency, as the single-lead and 12-lead signals were acquired almost simultaneously, with only a minimal temporal difference.

In contrast, the first approach relies on clinically provided reference values, but may involve comparisons between recordings taken on different days and under different conditions, which can introduce variability unrelated to the estimation method itself.
First approach:

Metric	NeuroKit		Our Model	
	MAE [ms]	Pearson *r*	MAE [ms]	Pearson *r*
PR	15.5	-0.9759	25	0.0211
QT	26.5	0.2861	16.3	0.7848
QTc	25.5	-0.9581	20.5	-0.2687

Table 4.3: Results of the comparison between CTIs obtained from single-lead signals processed with NeuroKit and our Model, with reference values automatically computed by the Mortara device and stored in the hospital 12-lead ECG database (in 4 patients).

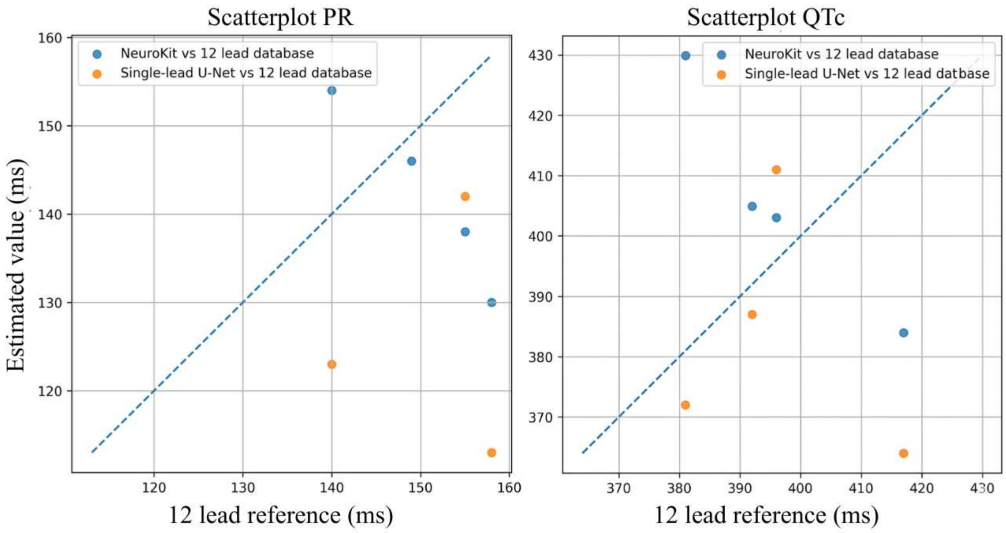

Figure 4.9: Scatterplots comparing PR and QTc intervals estimated from the single-lead ECG (NeuroKit and the proposed model) with reference values automatically computed by the Mortara device. Only three PR points are shown for the U-Net, as one signal had no detected P wave.

Second approach:

Metric	NeuroKit		Our Model	
	MAE [ms]	Pearson r	MAE [ms]	Pearson r
PR	24.5	-0.3721	14.3	-0.9922
QT	19.2	0.7924	4.0	0.9756
QTc	19.1	-0.3003	13.6	0.8556

Table 4.4: Results of the comparison between CTIs obtained from single-lead signals processed with NeuroKit and our Model against 12-lead signals used as input to the model (in 4 patients).

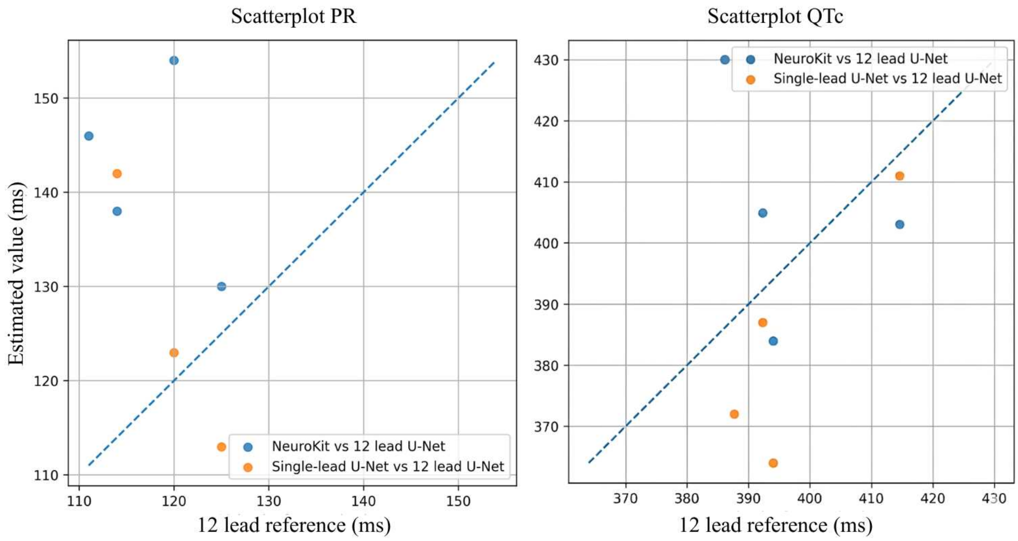

Figure 4.10: Scatterplots comparing PR and QTc intervals from the single-lead ECG (NeuroKit and the proposed model) with the 12-lead values obtained from our model. Only three PR points are shown for the U-Net, as one signal had no detected P wave.

As can be seen from the results, the second approach demonstrates greater stability. Using our model to estimate CTIs from single-lead signals produces values that are more representative and better correlated with the standard reference measurements obtained from the 12-lead ECG. For example, the MAE of our model is 14.3 ms for PR and 13.6 ms for QTc, compared to 24.5 ms for PR and 19.1 ms for QTc obtained using NeuroKit. Even a reduction of 10 ms is considered significant in this context. This improved performance is also confirmed by the Pearson correlation, which shows stronger agreement between single-lead and 12-lead measurements.

In contrast, the first approach shows less stability. For PR, NeuroKit seems to perform better, whereas for QTc, our model achieves a lower MAE, but NeuroKit shows a higher Pearson correlation. This indicates that relying solely on clinical reference values can lead to inconsistent comparisons depending on the interval considered.

Finally, in line with the reference CTI values automatically computed by the Mortara device at the hospital (first approach), the MAE and Pearson correlation were also evaluated across all patients, comparing NeuroKit and our model against the reference CTIs provided. For each patient, single-lead measurements were paired with the closest available 12-lead recording in time, rather than recordings acquired on the same day. This approach allows evaluating trends over a larger number of signals, but it may introduce variability unrelated to the estimation method itself due to temporal differences.

Metric	NeuroKit		Our Model	
	MAE [ms]	Pearson r	MAE [ms]	Pearson r
PR	24.4	0.539	30.78	0.205
QTc	31	0.344	19.97	0.467

Table 4.5: Comparison of CTIs obtained from single-lead signals processed with NeuroKit and our model, using reference values automatically computed by the Mortara device across all 40 patients.

The results indicate that performance depends on the interval considered. For PR, NeuroKit shows lower MAE and higher correlation, while for QTc, our single-lead method achieves lower MAE and higher correlation. Even when considering the values across a larger number of patients, the conclusions regarding the first approach remain unchanged. These findings highlight that no method consistently outperforms the other when using reference CTI values automatically computed by the Mortara device at the hospital.

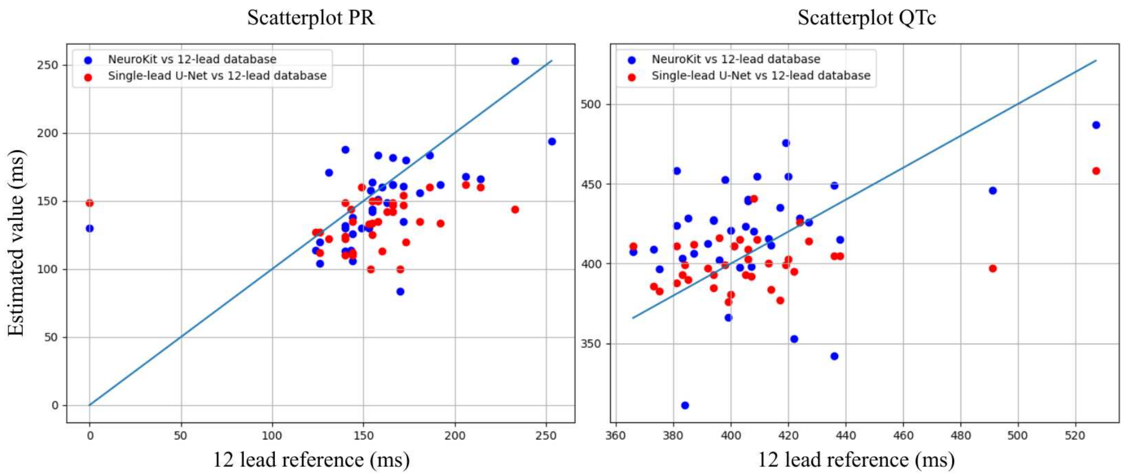

Figure 4.11: Scatterplots comparing PR and QTc intervals from the single-lead ECG (NeuroKit and the proposed model) with reference values automatically computed by the Mortara device, for all patients.

5 | Validation and Outcome Analysis

Based on Section 4.5.1, which shows that the proposed model outperforms Neurokit, we focus on the CTIs from the paired single- and 12-lead recordings, with the 12-lead ECGs acquired immediately after the corresponding single-lead signals, and perform a detailed analysis to draw conclusions.

This validation ensures that the single-lead CTIs from our model are reliable and comparable with the 12-lead reference. Building on this, a comprehensive analysis is then carried out to assess whether statistically significant differences emerge between the measurements. Both statistical and longitudinal approaches are applied to explore potential changes over time, such as increasing interval variability or the presence of shared trends across the analyzed subjects.

It is worth noting that predominantly negative findings are partly expected due to the imbalance of the dataset: only two out of forty patients experienced episodes of AF. Consequently, substantial variability or consistent trends across subjects are not anticipated.

5.1. Validation

In this section, the CTIs of interest (PR, QT, and QTc) were computed from both the single-lead and the 12-lead recordings acquired immediately afterwards, as in the second approach described in Section 4.5.1. This close temporal proximity ensures that the two signals are highly comparable.

To ensure accurate estimation of the PR and QTc intervals from the 12-lead ECG, specific leads were selected. For the PR interval, lead V1 was used, as the P wave, when present, is generally more clearly identifiable in this lead. The signal was processed using the same pipeline adopted for the single-lead recordings, namely by applying our trained model.

For the QTc interval, lead V5 was chosen to measure the QT interval. The correction was then performed using Fridericia's formula (Equation 2.1). In this case, the mean RR interval was computed across all leads of the corresponding 12-lead recording.

Patient	**PR (ms)**		**QT (ms)**		**QTc (ms)**	
	SL	12L	SL	12L	SL	12L
P1	113	125	342	356	372	387
P2	NaN	111	321	322	364	394
P3	142	114	379	380	387	392
P4	123	120	384	384	411	414

Table 5.1: Validation of CTIs between single-lead (SL) and 12-lead (12L) ECG signals acquired immediately afterwards, in four patients. PR, QT, and QTc intervals are reported in milliseconds (ms).

As shown in Table 5.1, the estimated CTIs from the single-lead recordings are consistent with those obtained from the 12-lead reference ECG. When the PR interval is reported as NaN, this indicates that the network was unable to detect any P-wave events in the signal.

The main limitation of this validation step is the small sample size. Although the agreement between the two recording modalities is satisfactory, the analysis was conducted on only four signals, which limits the strength of the conclusions that can be drawn.

5.2. Statistical and Longitudinal Analysis

Following the validation, a series of statistical and longitudinal analyses were performed on all the single-lead CTIs obtained from our model to investigate potential differences or temporal trends across patients.

5.2.1. CTI Temporal Trends

First, the temporal evolution of the PR and QTc intervals over the 90-day study period is presented for all patients.

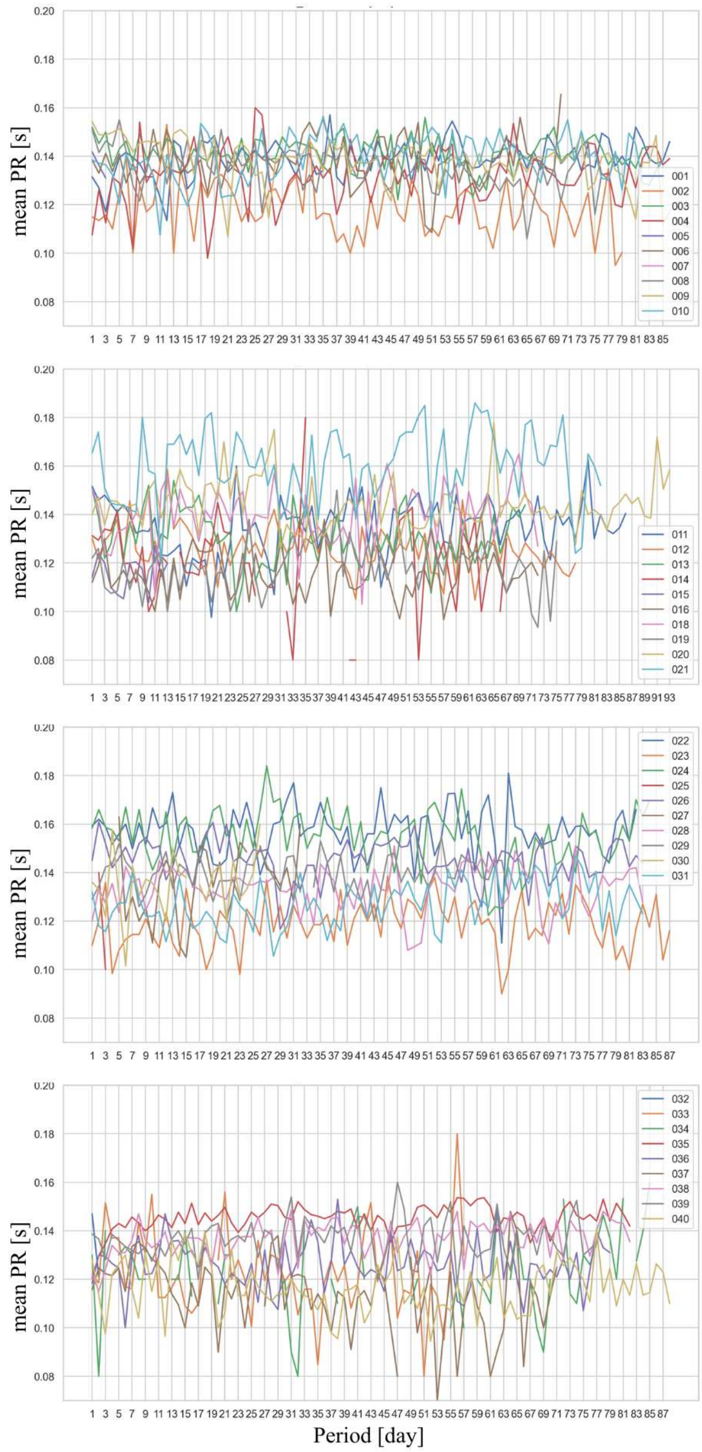

Figure 5.1: Trend of the PR interval throughout the study for all patients, grouped into sets of 10 per plot.

Figure 5.2: Trend of the QTc interval throughout the study for all patients, grouped into sets of 10 per plot.

The plots do not show any clear trends over the study period, with no consistent increases or decreases. PR intervals are uniformly distributed between 0.08–0.18 s, while QTc intervals range from 0.30–0.45 s. Because no obvious trends are visible qualitatively, we proceeded with quantitative analyses capable of detecting more subtle differences.

The first approach focused on evaluating beat-by-beat variability within each signal to determine whether significant fluctuations occur between consecutive beats. Specifically, for each signal, a linear regression was performed on the sequence of PR or QTc values across beats. The slope of the regression line indicates the trend (increase or decrease), and a t-test (significance level $\alpha = 0.05$) was applied to assess whether the slope was statistically significant. If the slope was significant and positive, the trend was classified as a significant increase; if significant and negative, as a significant decrease; otherwise, it was labeled as no significant change. Signals containing fewer than three beats were considered as insufficient data.

Beat-by-beat statistical significance analysis for PR and QTc intervals

Category	Percentage
Significant increase	2.8%
Significant decrease	2.7%
No significant change	79.7%
Insufficient data	14.8%

Table 5.2: PR beat-by-beat significance.

Category	Percentage
Significant increase	0.2%
Significant decrease	0.4%
No significant change	96.4%
Insufficient data	3.0%

Table 5.3: QTc beat-by-beat significance.

This beat-by-beat evaluation captures local variations in the intervals while providing a comprehensive overview of variability across the dataset. The percentages reported in Tables 5.2 and 5.3 summarize the proportion of signals showing significant increases, decreases, or no change in PR and QTc intervals between consecutive beats.

For the PR and QTc interval, the percentages of significant increases and decreases are roughly similar. This confirms that, overall, both PR and QTc intervals remain largely stable across beats.

After focusing on individual signals, and in particular on beat-by-beat variations, the analysis now shifts to a broader perspective by evaluating each signal as a whole.

The final analysis in this section assesses statistically significant changes between the beginning and the end of the study for each patient and each parameter using a t-test. Blue highlights indicate cases where a significant decrease was detected, red indicates a significant increase, and gray indicates no significant difference.

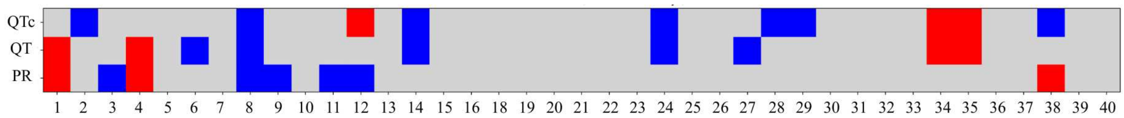

Figure 5.3: Heatmap showing significant changes in PR, QT, and QTc intervals for each patient during the study.

As expected, significant changes were not widespread across patients, confirming that, at the signal level, PR, QT, and QTc intervals remain generally stable over time.

5.2.2. CTIs and Quality of Life

The second major analysis focused on patients' quality of life. The hospital 12-lead ECG database, which contains CTI values automatically computed by the Mortara device from 12-lead ECG recordings, also includes several patient-reported quality-of-life variables.

The variables considered in this analysis were: Palpitations, Irregular heartbeat, Dizziness and fainting, Fatigue, and Shortness of breath. Each variable is measured on a categorical scale from 0 to 6, indicating how frequently the symptom occurred on a weekly or daily basis: 0 = never; 1 =

less than once per week; 2 = 1–2 times per week; 3 = 3 or more times per week, but not every day; 4 = at least once per day; 5 = several times per day; 6 = always. For this analysis, the goal was to correlate CTI changes over time with patients' reported quality of life. Only cases in which patients experienced a worsening in quality of life between the beginning and the end of the study were considered. Specifically, entries were selected when the symptom frequency reached a value of ≥ 2, indicating at least occasional weekly occurrence.

After identifying the relevant cases, pie charts were generated to illustrate the proportions of statistically significant increases, decreases, or non significant changes in CTIs between the start and the end of the study. Statistical significance was assessed using a t-test, which allowed the identification not only of significant differences but also of cases where no meaningful change occurred.

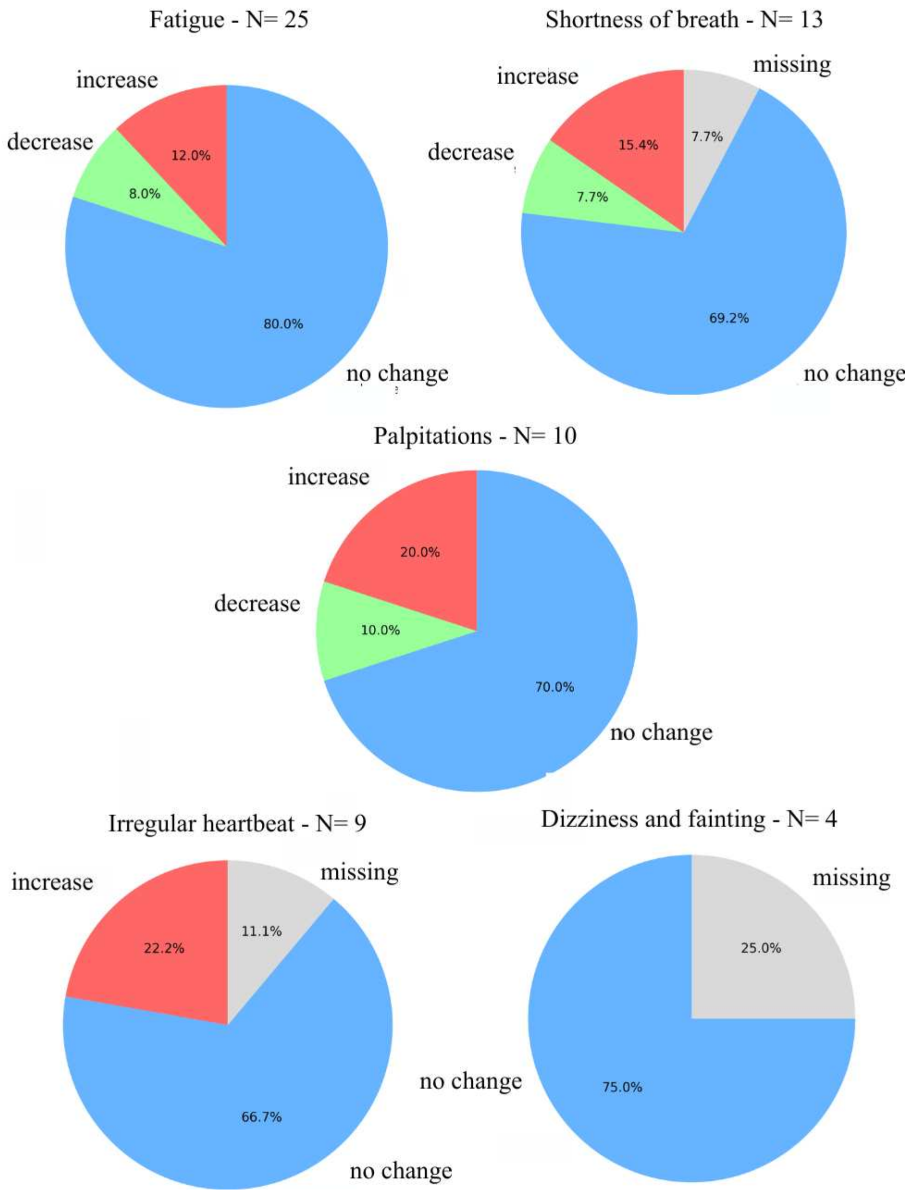

Figure 5.4: Each chart represents a quality-of-life variable according to significant PR interval changes.

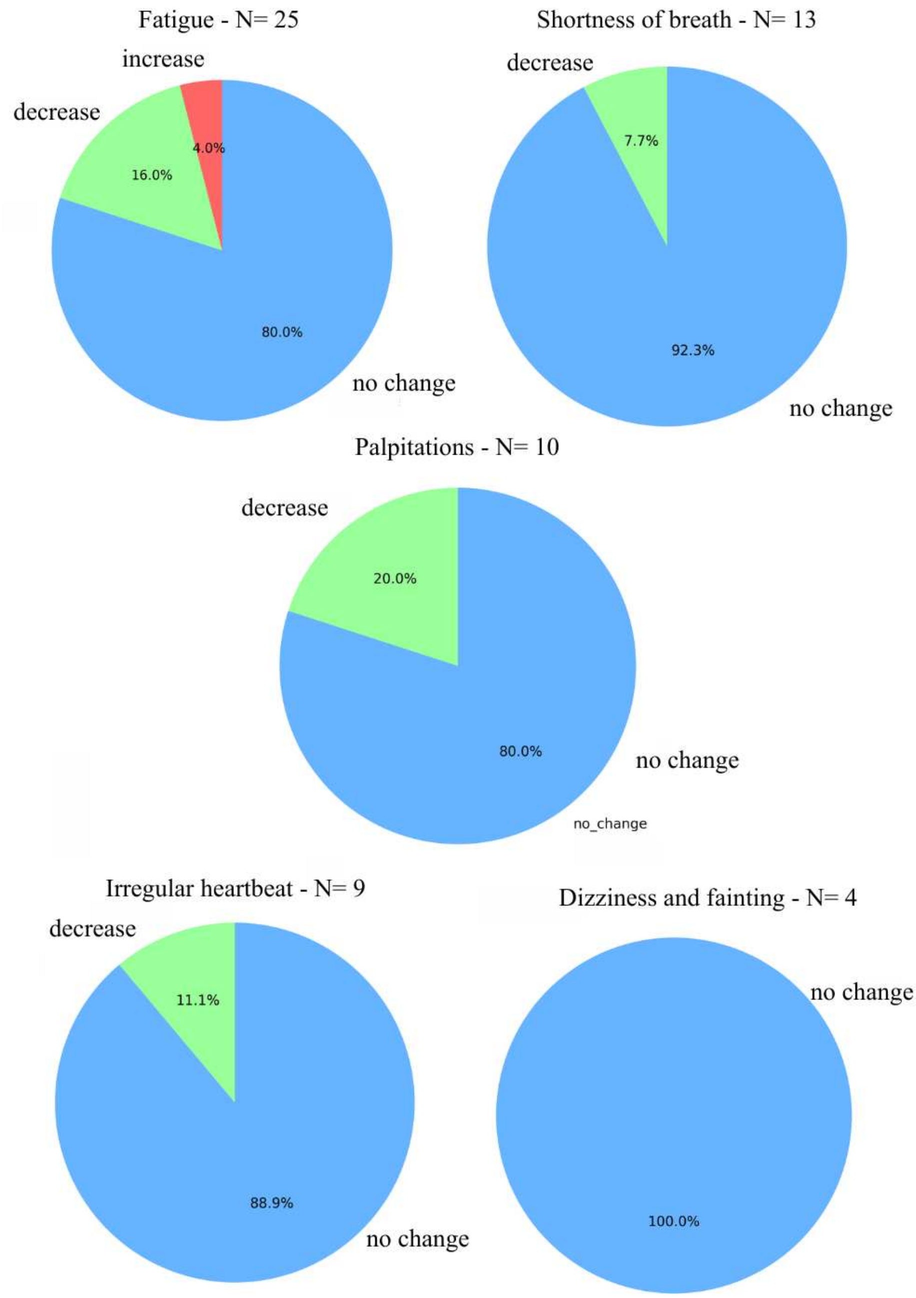

Figure 5.5: Each chart represents a quality-of-life variable according to significant QTc interval changes.

Each chart displays the percentage of patients within each subgroup who exhibited statistically significant differences in CTIs over the course of the study, grouped according to the categorical variable. “N” indicates the total number of patients reporting a worsening in quality of life for that specific variable. Overall, statistically significant changes in CTIs were not predominant among patients reporting a deterioration in quality of life. However, variables such as Fatigue, Shortness of breath, Palpitations, and Irregular heartbeats appear particularly relevant and may warrant careful monitoring in oncology patients. Missing values indicate cases in which the network was unable to detect P-wave events and, consequently, the PR interval could not be calculated.

5.2.3. CTIs and Oncological Treatment

The final analysis examined the oncological treatments administered to the patients, with the aim of identifying potential trends or differences in CTIs across treatment groups during the study period.

Drug	Therapy Type	N patients
Carboplatin + Pemetrexed	Cytotoxic chemo	2
Carboplatin + Vinorelbine	Cytotoxic chemo	8
Carboplatin + Etoposide	Cytotoxic chemo	2
Cisplatin + Vinorelbine	Cytotoxic chemo	2
Carboplatin + Pemetrexed + Pembrolizumab	Chemo-immunotherapy	5
Carboplatin + Paclitaxel + Pembrolizumab	Chemo-immunotherapy	1
Osimertinib	Targeted therapy	3
Lorlatinib	Targeted therapy	1
Carboplatin + Vinorelbine + Radiotherapy (RT)	Chemo-radiotherapy	1

Table 5.4: Summary of oncological therapies administered for lung cancer treatment.

The Table 5.4 reports the drugs administered and the type of therapy received by the patients. Not all patients are included in the counts, as treatment information was unavailable for a subset and was therefore classified as “NA”.

After defining the treatment groups, a cross-sectional analysis was performed first to assess whether statistically significant differences in PR and

QTc intervals existed among the therapies. The t-test did not reveal any significant differences across the groups. However, according to the literature, among all therapies included in this study, only Osimertinib has been reported to potentially induce QTc prolongation [42].

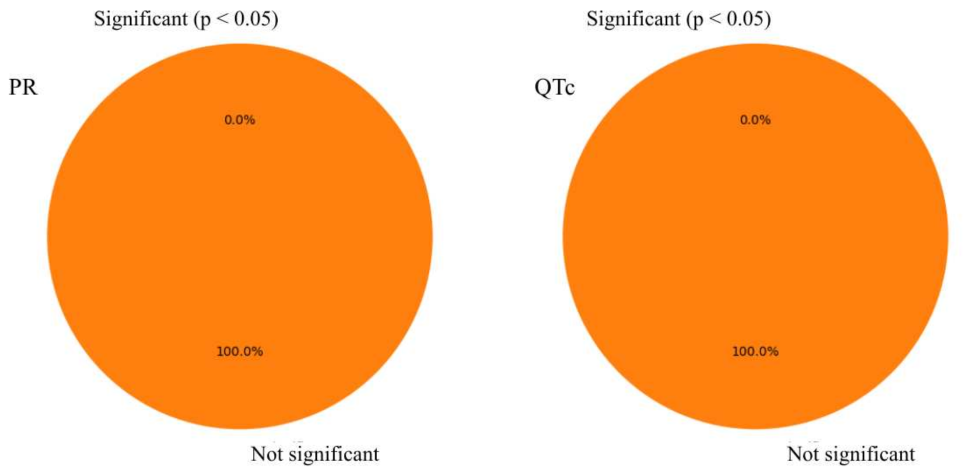

Figure 5.6: Pie charts of PR (left) and QTc (right) showing no significant longitudinal differences across therapy groups.

Continuing the analysis among treatment groups, the next step investigated parameter variability within each therapy group. The goal was to determine whether greater variability was observed in specific patients and to examine, within each therapy group, how statistically significant differences in PR and QTc intervals were distributed between the beginning and the end of the study.

To address the first aspect, boxplots were generated to represent the average variance of CTIs among patients in the same treatment group.

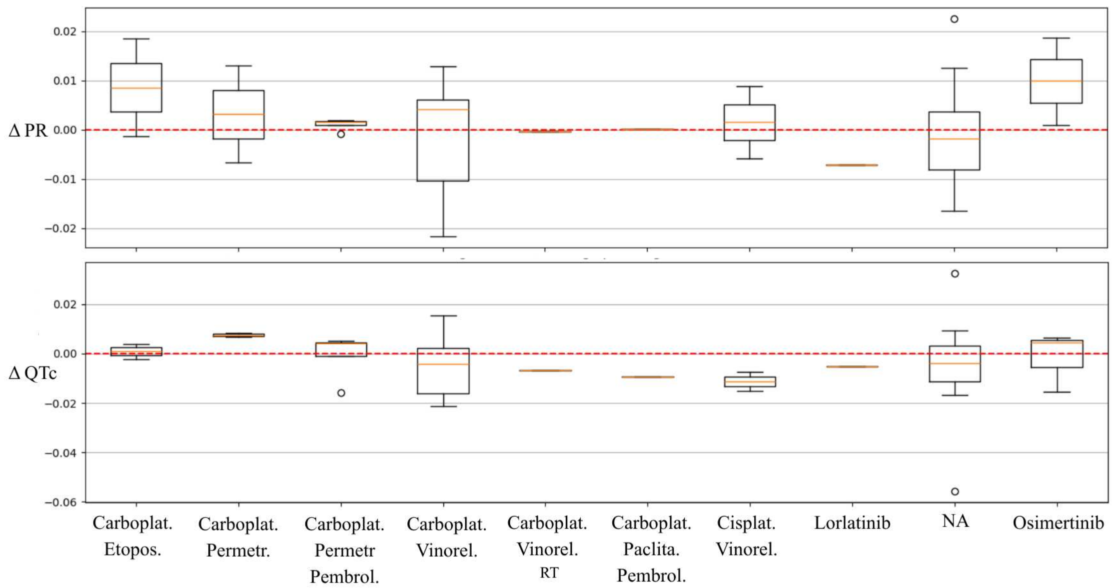

Figure 5.7: Boxplots showing CTI variability within each treatment group for both, PR and QTc.

No marked differences in variability were observed between treatment groups, although small variations may still be considered.

Finally, attention was focused within each treatment, and pie charts were generated to show, for each therapy, the proportion of patients exhibiting significant changes over the course of the study. Treatments for which no significant longitudinal differences were detected were excluded from this representation.

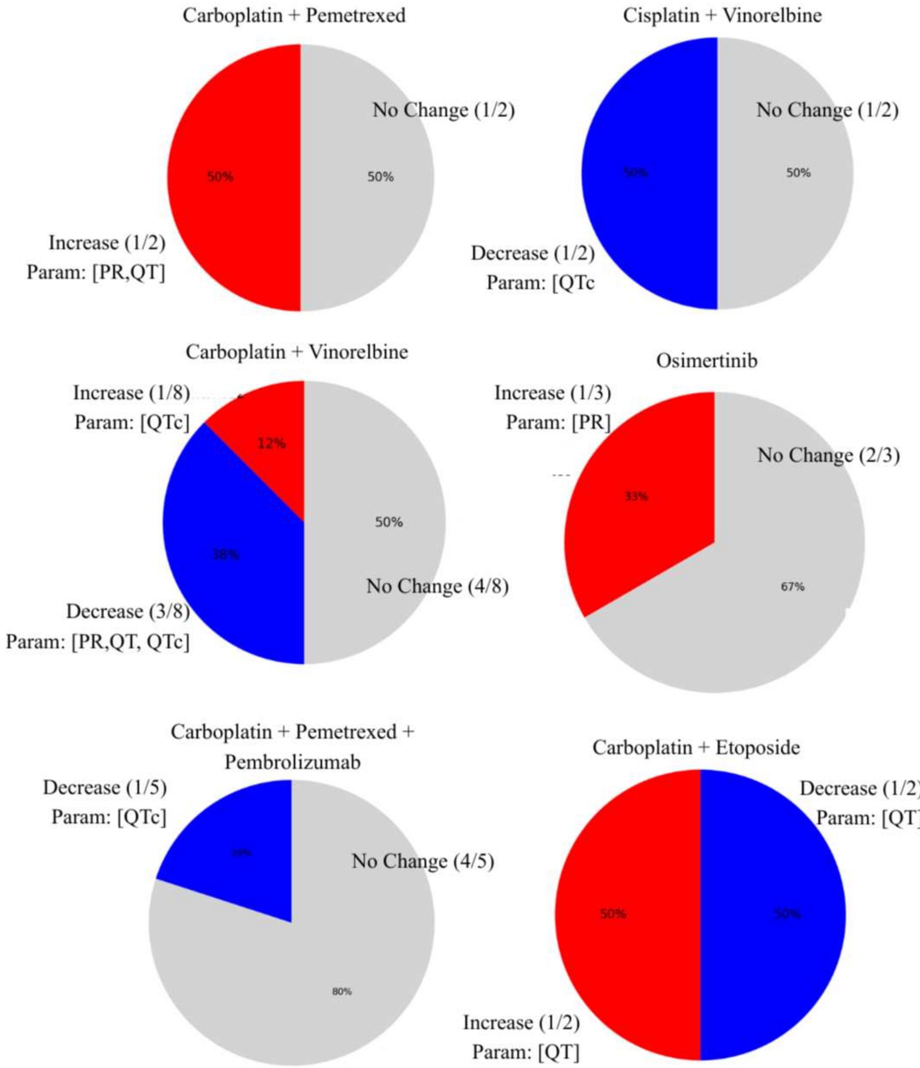

Figure 5.8: Distribution of significant longitudinal changes in PR and QTc within each treatment group.

As observed, Osimertinib did not show any impact on QTc in this context; no cases of significant QTc prolongation were recorded during the study for patients receiving this therapy. Other drugs showed some changes, highlighting both the percentages and the fractions of patients in each segment of the pie charts. The greatest variability was observed in the Carboplatin + Vinorelbine group.

6 | Discussion

This work demonstrates the feasibility of developing a neural network, followed by a post-processing phase, to segment an ECG signal acquired from a smartwatch. Unlike standard 12-lead ECGs, these signals are single-lead and generally lower in quality. To address this, a deep learning algorithm based on a CNN was designed and adapted to operate on one-dimensional signals, rather than the two-dimensional images for which CNNs were originally developed. By segmenting single-lead ECG signals, it became possible to reconstruct individual heartbeats and extract two key CTIs as final outputs: the PR interval and the corrected QT (QTc) interval. These parameters are clinically relevant, as they help assess potential variability and differences among patients, particularly in the context of AF. AF is a common cardiac arrhythmia, and early identification is crucial for timely intervention and prevention of severe complications.
In this study, AF was of particular interest because the patients were affected by lung cancer and underwent treatments such as chemotherapy, radiotherapy, or immunotherapy, which can lead to cardiovascular side effects including AF.

Although the dataset included only two patients who developed AF episodes during the study, the network demonstrated reliable performance in estimating PR and QTc intervals during the validation phase. This supports the potential clinical applicability of the proposed approach. Focusing on

CTIs offers a transparent and explainable framework, unlike many "black-box" models in the literature that detect or predict AF without providing clinically interpretable outputs.

Another important achievement of this work is the possibility of operating with low-quality signals acquired in real-world conditions. Single-lead smartwatch recordings allow continuous and long-term home monitoring without requiring frequent hospital visits. Integrating this approach into WDs could support 24-hour monitoring and facilitate early detection of cardiac abnormalities.

The literature confirms that few studies are closely aligned with clinical practice. Most focus on AF detection or prediction, while fewer directly estimate clinically meaningful intervals. QTc analysis is well established and widely investigated for AF risk, whereas the PR interval is less often used as a primary biomarker. However, the PR interval provides additional temporal information beyond P-wave duration, capturing atrioventricular conduction up to the onset of the QRS complex.

All analyses in this study were performed using the U-Net architecture, a CNN-based model widely applied in biomedical signal and image segmentation, including single-lead ECG analysis.

6.1. Limitations and Future Directions

While the study achieved promising results as a starting point for further investigation, several limitations must be acknowledged.

First, the limited number of patients who developed AF episodes represents a major constraint. As a consequence, no statistically significant correlation between chemotherapy treatment and AF occurrence could be established. Moreover, the validation phase was performed on only four signals, which further limits the generalizability of the findings. Overall, these limitations highlight the need for larger patient cohorts and more extensive validation datasets in future studies.

Another important limitation is the lack of sufficient 12-lead ECG recordings to perform a statistically significant comparison between baseline and follow-up measurements. In this study, only CTIs derived from a 12-lead ECG at baseline and after 90 days were available. Therefore, it was not possible to perform statistical analyses such as a paired t-test to confirm whether the differences observed in the single-lead signals were consistent with standard 12-lead measurements. A cross-validation study design, comparing single-lead and 12-lead ECG data over time, would strengthen the reliability of the findings.

Additionally, the temporal resolution of 20 ms used in the U-Net model limited the precision in estimating very short-duration intervals. Training the model with higher temporal resolution, although computationally more demanding, could improve its accuracy in detecting subtle variations in cardiac time intervals.

To address these limitations and further improve the proposed approach, several future developments can be considered:

- Expanding the dataset to include a larger number of patients and ECG recordings, particularly patients who develop AF episodes.
- Increasing the number of signals used in the validation phase to enhance statistical robustness.
- Refining and optimizing the U-Net architecture to improve temporal resolution and segmentation accuracy.
- Designing cross-validation studies comparing single-lead and 12-lead ECG measurements to strengthen clinical reliability.

These improvements could support more robust clinical analyses and enhance the interpretability and applicability of CTI-based biomarkers in real-world settings. Overall, this work demonstrates that CTI estimation from single-lead smartwatch ECGs is feasible and lays the foundation for future clinical studies with larger cohorts.

7 | Conclusions

This work demonstrates the feasibility of using DL to analyze low-quality, single-lead ECG signals acquired with WDs. By developing a U-Net based CNN adapted for one-dimensional signals, it was possible to segment individual heartbeats and accurately estimate key CTIs (PR and QTc), which are clinically relevant for monitoring AF. The study shows that even with limited and noisy data, meaningful physiological measurements can be extracted, bringing the analysis closer to clinical practice. Moreover, this approach supports continuous remote monitoring, allowing patients to be followed over time without frequent hospital visits.

In conclusion, the developed pipeline enables the analysis of low-quality ECG signals while providing clinically interpretable cardiac intervals. It opens the door to continuous, long-term monitoring and could serve as a foundation for detecting significant events and patient-specific variability.

Future work should focus on expanding the validation dataset, integrating remote monitoring solutions, and exploring predictive models for early detection of AF, particularly in oncology patients undergoing potentially cardiotoxic treatments. These improvements would enhance both the reliability and the clinical applicability of the proposed approach.

8 | Appendix

8.1. Literature review search strategy

This section describes the strategy and search queries used for the state-of-the-art review, developed according to the PRISMA guidelines [43]. For this analysis, the databases explored are PubMed, IEEE Xplore, and Scopus, selected for their relevance in the scientific research domain.

8.1.1. Keywords

Based on the research topic, the keywords selected for the analysis are:
"AF Prediction" OR "Atrial Fibrillation Prediction"
"Single Lead ECG" OR "Single-Lead ECG"
"Lung Cancer" OR "Pulmonary Neoplasms"
"Chemotherapy"
"CTI" OR "Cardiac Time Interval"
"STI" OR "Systolic Time Interval"
"QT-Interval"
These keywords were used to create search queries, which are illustrated in the following paragraph.

8.1.2. Queries

Based on the keywords reported in the previous section, the queries used to develop this chapter were created through an iterative process.

8| Appendix

After several attempts, it was possible to define queries that retrieve the most relevant papers while minimizing irrelevant results.
The initial queries returned a very large number of results, many of which were unrelated to the topic of interest.
To address this, the keywords were carefully combined using Boolean operators to narrow the search to the specific problem addressed.
The final queries are:
Query 1: (“CTI” OR “CARDIAC TIME INTERVAL” OR “STI” OR “SYSTOLIC TIME INTERVAL”) AND ("SINGLE LEAD ECG" OR "SINGLE-LEAD ECG")
Query 2: "QT INTERVAL" AND ("SINGLE LEAD ECG" OR "SINGLE-LEAD ECG")
These queries allow the literature review to be structured in two main phases:

- Phase 1: A general analysis of methods applicable to single-lead ECG signals for AF classification. This represents a classification-based approach rather than a prediction one, focusing on the analysis of cardiac intervals and exploring what can be implemented for studying atrial fibrillation.
- Phase 2: A clinically oriented approach, aimed at understanding which aspects physicians typically analyze in a single-lead ECG to determine a higher probability of AF, and how the relevant features

can be automatically extracted for such analysis.

8.1.3. Selection Process

Once the queries were defined, their application to the previously mentioned scientific databases produced a series of results:

Database	Query 1 results	Query 2 results
Pubmed	0	17
IEEE Xplore	31	9
Scopues	64	65

Table 8.1: Results of query 1 and 2 across selected database

The papers obtained were then filtered and selected according to the following procedure:

- Time frame: the analysis considered studies published within the last decade. Therefore, all articles published between 2014 and 2025 were selected for further review.
- Screening process: the papers were analyzed in multiple stages. First, the title was examined, and any paper not relevant to the research topic was discarded. Next, the abstract was reviewed, and only those papers consistent with the research objectives were selected for full-

text analysis.

8.2. Review of Atrial Fibrillation's Detection

In recent years, many new approaches have been developed for the detection of AF, with the aim of improving patient health. Thanks to the rapid development of digital technologies, in particular artificial intelligence, machine learning and deep learning, these methods have become increasingly effective for medical signal analysis. Consequently, numerous studies have contributed to the expansion of both scientific and clinical knowledge in this field. In this part of literature review, the focus is on understanding how previous studies have achieved AF detection and on comparing and analyzing the strategies used to build reliable models for classifying AF and a wide range of arrhythmias. It is important to note that the selected articles deal with the detection, rather than prediction, of AF.

The most relevant aspects that differentiate these methods, and allow for meaningful comparison, are: the dataset used for the investigation, the computational approach adopted (ML, DL or statistical methods) and the features extracted from the ECG data for AF detection. Nevertheless, most of the reviewed studies share a common goal: to classify ECG signals into four categories. The algorithms developed in these studies are typically designed to distinguish among Normal Sinus Rhythm (NSR), AF, Noisy signals, and Other rhythms (such as bradycardia or tachycardia).

At the end of this section, **Table 8.2** summaries the different methods discussed, based on the aspects mentioned above, which will be further explored in the following paragraphs.

8.2.1. Implemented Methods

All the reviewed studies follow a similar three-step process:

> 1) Signal processing: this step usually involves applying a Butterworth filter to remove noise at both low and high frequencies, followed by downsampling to reduce the computational cost of the algorithm.
> 2) Signal transformation: the processed signal is either used for feature extraction or transformed into an image, converting the data from the 1D to the 2D domain.
> 3) Classification: finally, the selected classifier is applied to perform the classification task.

Some authors have investigated AF from single-lead ECG signals using ML or DL approaches. Among these, DL has become the dominant approach, although some studies have also employed ML methods. Regarding ML-based approaches, among the studies that implemented ML algorithms [44], [45], [9], two of them [44],[45] employed a Random Forest classifier. The Random Forest algorithm constructs multiple decision trees on different subsets of data. The key condition is that these subsets must have minimal correlation with each other. According to the bagging theory, several uncorrelated models can provide more accurate predictions than any individual model. This is because the models "protect" each other from their individual errors, as long as they do not consistently fail under the same conditions. The main advantages of the RF algorithm include its ability to efficiently process large datasets, its robustness to feature scaling, its effectiveness in reducing overfitting, and its usefulness for feature selection. Yavorskyi et al., 2021 [45] and Mahajan et al., 2017 [44] are very similar studies. Both used the PhysioNet/Computing in Cardiology 2017 dataset (accessibility: public), which contains approximately 8.5K single-lead ECG recordings, and applied RF algorithm. In the signal processing stage, both [44] and [45] detected R-peaks using the Pan–Tompkins algorithm [46], obtained RR intervals, and extracted several features from them. In [44], linear, non-linear, frequency, and time-domain methods were applied to extract features from different domains, in order to characterize the signals by their morphology and temporal duration. Furthermore, to improve the generalizability of the classification model, redundant features were removed using a Genetic Algorithm-based feature dimension reduction technique, reducing the feature set from 62 to 37.

In contrast, [45] selected 22 features, including the average value of the RR interval duration (NN), the standard deviation of all RR intervals (SDNN), and several others. After training the models, the overall F1-scores obtained with the proposed classification models were 0.74 in [44] and 0.959 in [45] for AF detection on the test dataset.

Among ensemble methods, the study by Bogatinovski et al. 2017 [9] is also noteworthy, where the algorithms AdaBoost and Gradient Boosting were implemented.

	Class	N	SVEB	VEB	F	Q
	N	43832	25	396	3	0
True	SVEB	1721	15	100	1	0
Label	VEB	708	76	2432	4	0
	F	281	1	106	0	0
	Q	3	0	0	4	0

Figure 8.1: Confusion matrix for the best performing model Adaboost, number of trees 500, learning rate 0.1 [9]

In this case, the aim is not to classify AF directly, but to classify five different types of heartbeats: N (non-ectopic), SVEB (supraventricular ectopic beat), VEB (ventricular ectopic beat), F (fusion beat), and Q (unknown beat). Starting from a single-lead dataset, the authors do not extract local statistics from the ECG to use as input features for the network. Instead, they focus on global characteristics in order to achieve a better description and classification of heartbeats. The HCTSA (Highly Comparative Time Series Analysis) library is used to extract global features from the time series. These features originate from interdisciplinary studies in fields such as statistics, economics, physics, and biomedical signal processing, and they quantify patterns over the entire time interval. After extracting the heartbeats and applying the HCTSA library, the most relevant features are selected using a Random Forest. Based on the impurity score of each feature, only those exceeding a relevance threshold are retained. The selected features are then used as input for AdaBoost, which produces the final classification. While Random Forest is widely used in ML, CNNs represent the most commonly employed algorithm in the DL field. CNNs are deep learning architectures typically applied to inputs with a grid-like structure, such as images. Due to their sparse connectivity and parameter sharing characteristics, they allow to automatically discover complex and high-level representation from the images by using a well-trained network to detect AF from a single-lead ECG. In our case, the first step, after preprocessing, is to transform the 1D single-lead ECG signal into a 2D image. The studies [10], [3], [5], [11] have used CNN in order to classify AF:

- Fang et al., 2023 [10] and Lai et al., 2019 [3] share the idea of training two CNNs in parallel, which distinguish between AF and other signals, and also the idea of combining the two final results to achieve better overall performance. The differences between the two implementations are shown in the following figures, particularly regarding the input channels. Study [10] uses as input channels Poincare plot for the first channel and Spectrogram for the second channel. In contrast, study [3] uses as input channels RR interval series for the first channel and F-wave frequency spectrum for the second channel.

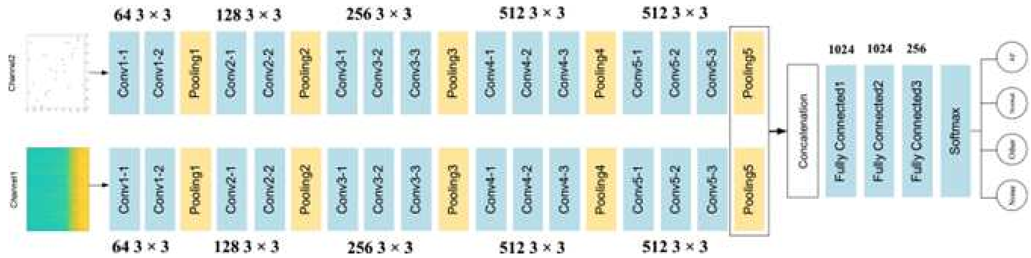

Figure 8.2: The architecture of the proposed dual-channel deep neural network. Input channel one is Poincare plot, and for channel two is Spectrograms [10]

For [3] is already shown in Figure 3.1

After preprocessing, [10] extracts time-frequency Spectrograms and Poincare plot both of which are used as input for the dual-channel neural network. The spectrogram reflects characteristics in the time domain, while the Poincare plot shows the variation of RR intervals based on the degree of point aggregation. The dataset, obtained from the 2017 PhysioNet/CinC Challenge, is divided in an 8:1:1 ratio to create training, testing and validation sets. The final accuracy on the test set is 0.87, and the overall F1-score is 0.83. In this project, the proposed dual-channel deep neural network could experience overfitting after approximately 27 epochs, so an early stopping technique is applied to address this issue. In contrast with existing methods that use deeper architectures, [3] presents an approach employing a CNN with representative rhythm features of AF rather than raw ECG signals, without any electrophysiological assumptions. Specifically, this work extracts RR interval series and the F-wave frequency spectrum as input to dual lightweight channels. Although the network is not very deep, it achieves a final accuracy of 97.5%.

- Alam R et al., 2023 [5] propose an alternative application of CNN. The objective is not classification, but the estimation of QT intervals and heart rates using a single-lead ECG, from 12-lead ECG dataset. The proposed method is a residual neural network, QTNet, with the architecture is already shown in Figure 3.2. Hence, the contributions of this work are twofold. First, the development of a new DNN model for regression of QT intervals from raw Lead-I ECG, and second, the evaluation of the proposed model on external datasets, including publicly available resources. QTNet demonstrates robust regression performance across all four test sets. This work presents a novel regression approach and shows a high correlation between the predicted values and the corresponding 12-lead labels across four independent datasets. Across all four datasets, the mean absolute error (MAE) in the estimated QT interval ranges between 9 ms and 15.8 ms. Pearson correlation coefficients vary between 0.899 and 0.914. In contrast, QT interval estimation on these datasets using a standard method for automated ECG analysis (NeuroKit2 [36]) yields MAEs between 22.29 ms and 90.79 ms, and Pearson correlation coefficients between 0.345 and 0.620. Estimation of QT intervals is useful for AF detection, as will be discussed in the next paragraph.

- The last study using a DL algorithm is Liu et al., 2019 [11]. The approach involves constructing an ensemble deep learning architecture for AF detection, where 1D signals and 2D images are used as inputs to the component learners. The signals are converted into time-frequency spectrums and Poincare plots.

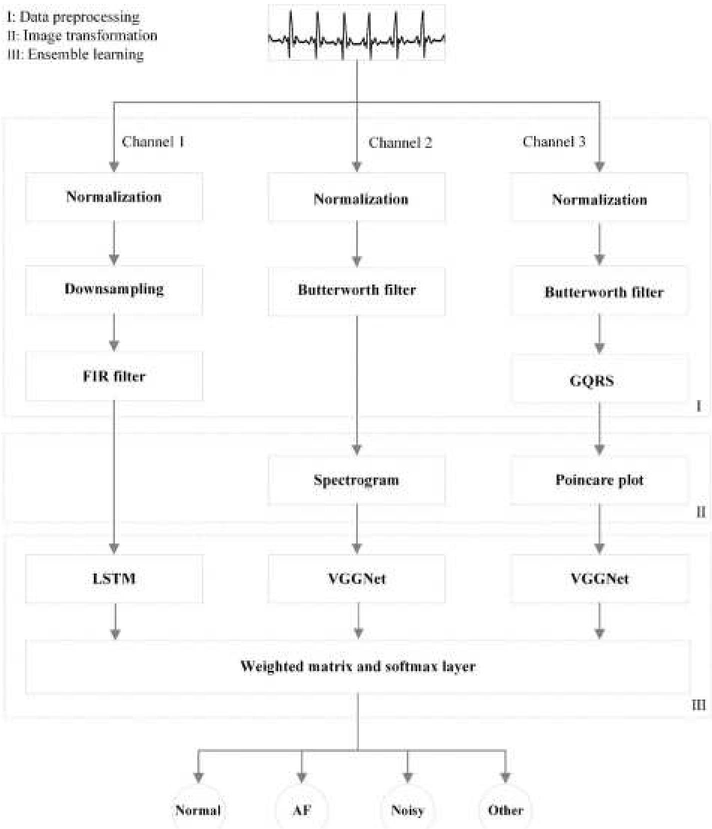

Figure 8.3: Diagram of the proposed model with three modules, data preprocessing, image trasformation and ensemble learning; where VGGNet is a typology of CNN [11]

Each learner produces its individual confusion matrix and F1 score. To

generate an accurate final output, a weighted matrix is constructed, and a probability value is obtained: $P = \sum_{i=1}^{3}(\Omega_i \cdot P_i)$, where i denotes the index of the component learner, Ω_i refers to the 4×4 weighted matrix of the i-th component learner, Pi represents 4×1 probability matrix of the i-th component learner, respectively.
The production $P = \{p_{\text{AF}}, p_{\text{Noisy}}, p_{\text{Normal}}, p_{\text{Other}}\}$ represents the calculated classification probabilities for the ECG signal. The final output is the maximum of these four values, which determines the classification result. The overall F1-score is 0.82, and the accuracy is 0.9, on test set.
The last method developed using a DL approach is by Soliński et al., 2017 [47]. The first step is the extraction of RR intervals from the QRS complex using a hybrid approach with two complementary methods applied hierarchically. The first method consists of a nonlinear transformation and a first-order Gaussian differentiator. If it successfully extracts the RR intervals, the workflow continues; otherwise, the second method, the Pan–Tompkins algorithm, is applied. After the hybrid detection step, the method developed the noise detection step in order to remove signals with high noise or lacking physiological from the main classification path. Noise detection is based on four extracted features: a signal is classified as "too noisy" and is not evaluated in the subsequent steps of the algorithm if at least one feature indicates a high level of noise. The next step involves basic rhythm classification to identify rhythms other than sinus and AF in the most evident cases. This is done by analyzing RR intervals using the

heart rate and pNN50 (percent of decelerations between two consecutive RR intervals grater than 50 ms), and by searching for characteristic patterns indicating ectopic beats, bigeminy, or trigeminy. A signal is classified as an “other” rhythm when the heart rate and pNN50 exceed threshold values or when the appropriate number of ectopic patterns is detected. After this initial filtering, 78 features are extracted from the signal, which are then used to train a multilayer perceptron (MLP) with two hidden layers.

• The hybrid QRS detection improves the reliability of R-peak identification, leading to better RR features and improved discrimination between AF and non-AF.

• Discarding strongly noisy signals and classifying obvious patterns reduces the load on the ML classifier and improves overall accuracy.

• The final result achieves an overall F1-score of 0.77.

All the approaches described are interesting for AF detection, but better results can be achieved by combining ML and DL methods, as demonstrated by Li et al., 2025 [12]. In this study, five algorithms were developed to estimate QT intervals. The five state-of-the-art QT interval estimation algorithms include one open-source signal processing-based method and four commercial machine learning-based algorithms. The results of these algorithms are combined in two ways: by taking the median or by using a linear regression model, in order to achieve higher accuracy.

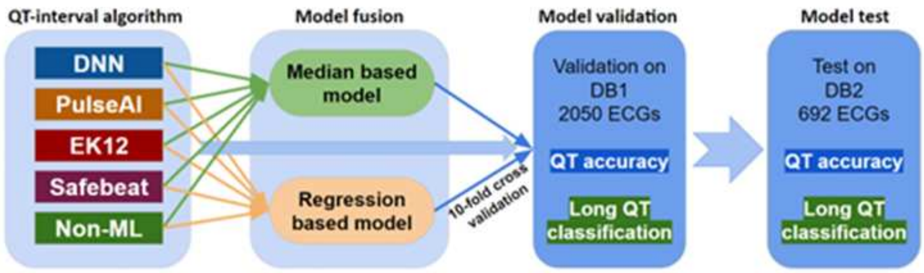

Figure 8.4: The flowchart of the algorithm [12].

In this study, the signals are not single-lead; nevertheless, it is an important and innovative work. The study demonstrates that the fusion of five independently developed, state-of-the-art QT estimation methods, based on very different approaches, significantly improves automated QT analysis and long QT detection using a 2-lead mobile ECG.

In addition, the investigation of AF was not limited to ML and DL approaches, but it also included statistical analysis. Indeed, studies [48] and [6] attempt to classify NSR and AF from single-lead ECG by examining the fundamental statistical differences of certain features between subjects. The difference between [48] and [6] lies in the method of feature extraction, as follows:
- Study [48] combines Wavelet Transform, which allows analysis of a signal in both time and frequency domains, with Higher-Order Statistics, mathematical techniques that capture non-linearities and relationships between multiple frequencies. The combination of these methods, referred

to as WHOS (Wavelet Transform + Higher-Order Statistics), produces the Wavelet Bispectrum (WBS), an advanced representation of the ECG that reveals Quadratic Phase Couplings (QPC), interactions between pairs of frequencies that generate new components, and temporal evolution, showing how these interactions change over successive heartbeats. Using this approach, 50 features were extracted, including local, global, and temporal dynamic features.
- Study [6] extracts several features based on P-waves and RR intervals. Among these, the most discriminative for identifying susceptibility to AF were the P-wave duration, the standard deviation of the beat-to-beat Euclidean distance (expressed as mean and standard deviation) between consecutive P-waves, and the sample entropy of the RR intervals.

The statistical comparison between the two groups, in both [48] and [6], was performed by using the Mann-Whitney U test for unpaired data. Results were considered to be statistically significant at $p<0.01$. In [6], the patients with a history of AF presented significantly longer P-wave duration and higher mean and standard variation of the beat-to-beat Euclidean distance when compared to the control group. On the other hand, HRV analysis showed that patients with a history of AF had lower sample entropy than patients without previous AF history, meaning that RR intervals exhibited lower complexity in subjects with AF. To assess the robustness and generalizability of the classifier, a 10-fold cross-validation was applied, resulting in a global accuracy of 86%. Instead in [48], the statistical comparison between NSR and AF time series features revealed significant differences, indicating non-linear and chaotic characteristics in the evolution of AF WHOS dynamics.

8.2.2. Clinically Useful Biomarkers for AF

Once an overview of the methods used for AF detection has been provided, it is now interesting to focus on a more clinically oriented aspect. This means highlighting the biomarkers that are mainly used in clinical practice by physicians to investigate AF. First, the RR interval is derived, from which heart rate can be easily obtained. To compute RR intervals, it is first necessary to identify the individual R peaks. Once the R peaks and their sequential arrangement have been detected, it becomes possible to analyze features related to the cardiac cycle. In the previous study, several features were shown to be correlated with AF, such as the QT interval, P and F waves. Let examine them in detail. As reported in [3], RR intervals and F-waves are characteristics directly correlated with AF. AF is manifested on the ECG signals as the absence of P-waves and the presence of fibrillatory waves (F-waves – see Figure 2.4). Another important feature is an irregular ventricular rate, which represents the variability of RR intervals [30], [49], [50].

From Figure 2.4 is possible to notice a clear variability in R-R intervals between heartbeats, with some intervals being much longer than others. In AF signals, the P-wave is often absent, while numerous low-amplitude F-waves appear where the P-wave would normally be. AF can also be observed through the characteristics of the P-wave, as reported in [6]. The duration of the P-wave reflects abnormal conduction in the atria, and P-wave indices such as maximum duration, dispersion, area, or axis can be used to identify patients at increased risk of developing AF. The representative features include P-wave duration, the mean and standard deviation of the beat-to-beat Euclidean distance, and the sample entropy of the RR intervals.

	Sensitivity	Specificity	Accuracy
F1	55%	95%	76%
F2	41%	95%	70%
F3	63%	75%	70%
F1 + F2	61%	92%	78%
F1 + F3	63%	92%	78%
F1 + F2 + F3	81%	92%	86%

F1: P-wave duration.
F2: Standard deviation of beat-to-beat Euclidean Distance.
F3: Sample Entropy.

Figure 8.5: Classification results [6].

The last opportunity to study AF detection is related to the QT segment. After detecting the Q and T points, the QT interval can be analyzed. The QT interval, corrected for RR interval (QTc), for example using Fridericia's formula, is important because QT prolongation, as noted in [5], often leads to fatal arrhythmias and sudden cardiac death. Antiarrhythmic drugs can increase the risk of QT prolongation, making post-treatment monitoring and dosage control essential.

8.2.3. Automatic Feature Extraction from Data

In this section, we analyze how to automatically extract clinically relevant features. The first fundamental step is the detection of R peaks, which can be performed using well-established methods such as NeuroKit [36] or the Pan–Tompkins algorithm [46] to directly obtain the RR interval. Once the R peaks have been detected, a search window preceding the R-wave can be applied to identify the P-wave [6]. The width and position of this window can be fixed for all beats but may vary depending on the subject. The onsets and offsets of the P-waves can be determined using first- and second-derivative approximations of the ECG signal. It is also useful to compute the beat-to-beat Euclidean distance between P-waves as a measure of P-wave morphological variability over time. The P-wave duration is calculated as the difference between its offset and onset, while the beat-to-beat Euclidean distance is computed according to the formula described in [6].

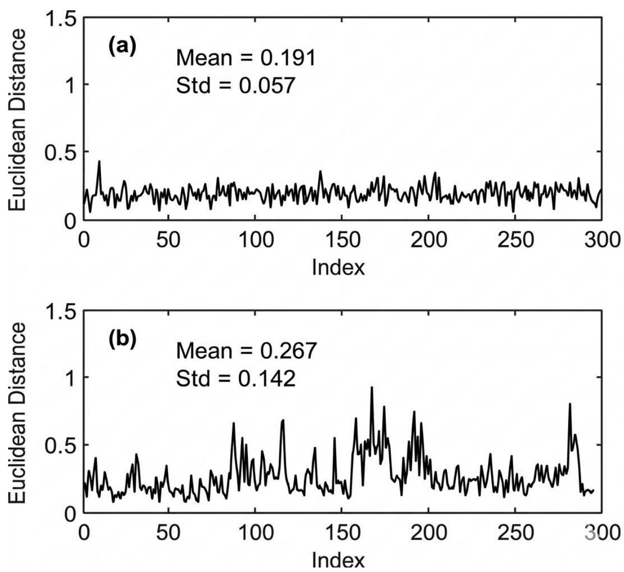

Figure 8.6: Beat-to-beat Euclidean distance of extracted P-waves for a healthy subject (a) and a patient with AF history (b) [6].

Alternatively, instead of analyzing the P-wave, it is possible to investigate the F-wave [3], [51]. These are rapid and irregular oscillations that occur within a specific frequency range, typically between 2 and 9 Hz. For this reason, they can be identified through spectral analysis methods, such as the Fourier Transform or the Wavelet Transform.

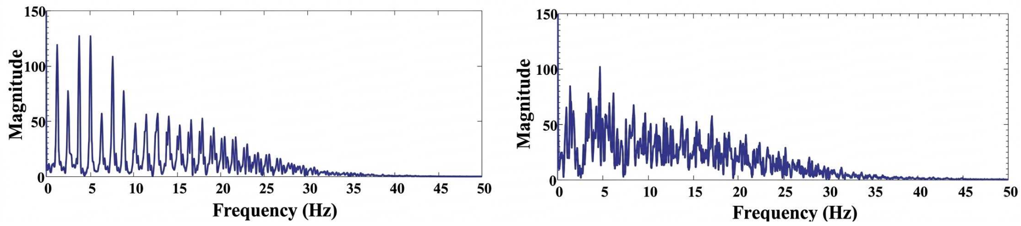

Figure 8.7: Feature extraction of cardiac rhythms both of normal subject (left) and patient with AF (right): F-wave frequency spectrum [3].

This spectrum highlights the presence of high-frequency oscillations in the ECG signal, which are indicative of AF. The last automatic detection method concerns the calculation of the QT interval, which can be obtained as described in [5]. A residual CNN adapted for one-dimensional ECG signals is presented. The network is designed to simultaneously estimate the QT interval and heart rate from 10-second ECG segments. This is possible because the final layers of the model are a multilayer perceptron that performs regression on these two parameters.

Table 8.2: Comparison of studies on AF

Author	**Dataset**	**Method**	**Label**	**Results**
R. Mahajan et al., 2017 [44]	8528 single-lead ECG, PhysioNet/ CinC 2017	ML: random forest	Three-class classification: Normal, AF, and Other	Overall F1-scores: 0.78 on the test dataset

8| Appendix

A.Yavorskyi et al., 2021 [45]	8528 single-lead ECG, PhysioNet/ CinC 2017	ML: random forest	Four-class classification: Normal, AF, Noisy and Other	Overall F1-scores: 0.956 on the test dataset
Bogatinovski et al., 2017 [9]	MIT-BIH arrhythmia database. N=47 subjects. Available on Physionet	ML: RF for relevant feature selection, adaboost/ gradient boosting as classifier	Five-class classification, not AF	Features from different domains help separate classes in heartbeat classification
Fang B et al., 2023 [10]	8528 single-lead ECG, PhysioNet/ CinC 2017	DL: dual channel neural network	Four-class classification: Normal, AF, Noisy and Other	The final accuracy on the test set is 0.87, and the overall F1-score is 0.83
D. Lai et al., 2019 [3]	AFDB database. Accessibility: public, in PhysioNet	DL: dual channel lightweight CNN	Binary classification: AF vs. Not AF (Normal Rhythm)	Final accuracy of 97.5%
Alam R et al., 2023 [5]	4,22 million of ECG signals, from 903k patients. Accessibility: private	DL: CNN + residual blocks for QT and HR regression	Regression problem, not label	MAE in the estimated QT interval ranges between 9 ms and 15.8 ms. Pearson correlation coefficients vary between 0.899 and 0.914

8| Appendix

Y. Liu et al., 2019 [11]	8528 single-lead ECG, PhysioNet/ CinC Challenge 2017	DL: ensemble deep learning with three channels	Three-class classification: Normal, AF, and Other	The overall F1-score is 0.82, and the accuracy is 0.9 on the test set
M. Soliński et al., 2017 [47]	8528 single-lead ECG, PhysioNet/ CinC Challenge 2017	ML approach + hybrid QRS detection	Three-class classification: Normal, AF, and Other	Overall F1-score of 0.77
Qiao Li et al., 2025 [12]	Many databases used; some public, some private (2-lead mobile ECG)	ML + DL algorithms	Estimation of QT and QTc intervals	The fusion of five state-of-the-art QT estimation methods improves automated QT analysis
C. A. Zisou et al., 2022 [48]	5834 single-lead ECG. Accessibility: public, available on Physionet	Statistical analysis with Wavelet Transform + Higher-Order Statistics	Not classification	Mann–Whitney U tests revealed significant differences between NSR and AF
A. Luca et al., 2016 [6]	N = 76 patients. Accessibility not mentioned	Statistical analysis	Not classification	Group differences were assessed using the Mann–Whitney U test, and 10-fold cross-validation achieved an overall accuracy of 86%

8.3. Discussion on Research

Analyzing the obtained results, it appears that the most commonly used methods for studying AF belong to the DL family, in particular CNN architectures. However, ML methods and statistical analysis approaches are also used and remain relevant.

The DL method that shows the highest accuracy is presented in [3]. This study demonstrates that representative ECG features can effectively capture AF rhythm characteristics and can be successfully used as inputs to CNN architectures, achieving high classification performance while keeping the network relatively simple. Despite these promising results, some limitations are still present. One important limitation is the use of a small dataset (25 signals). Moreover, the reliance on specific feature representations, such as the F-wave frequency spectrum, may limit the applicability of the approach in datasets where AF episodes are not clearly present. Additionally, the limited dataset may increase the risk of overfitting, potentially reducing the generalization capability of the model in real-world scenarios.

Alam R et al., 2023 [5] is relevant from both technical and clinical perspectives. The study addresses QT interval regression, a clinically meaningful parameter in AF research, demonstrating the feasibility of adapting CNN architectures to estimate ECG intervals directly from single-lead signals. However, the approach remains largely “black-box” in nature, limiting interpretability from a clinical standpoint. Another limitation is that the

single-lead signal is extracted from a 12-lead ECG dataset, implying higher signal quality compared to typical single-lead acquisitions. Additionally, the study highlights the widespread use of NeuroKit2 as a benchmark for ECG interval estimation.
Finally, A. Luca et al., 2016 [6] provide clinically relevant insights through a statistical analysis of ECG features. The study emphasizes the importance of P-wave characteristics and RR interval variability in identifying AF susceptibility. In particular, the results highlight the relevance of P-wave duration and entropy-based RR features as discriminative biomarkers. However, statistical approaches may be limited in capturing complex non-linear patterns in ECG signals compared to modern DL-based methods.
Overall, the reviewed studies highlight existing gaps that this work aims to address: the need to operate effectively with scarce data, to develop more transparent and clinically interpretable approaches, particularly through the analysis of clinically relevant CTIs, and to enable a home-monitoring framework based on single-lead ECG signals.

Bibliography

[1] Nagwa. Understanding the human circulatory system, 2025. URL https://www.nagwa.com/en/explainers/912123271719/. Accessed: 2025-03-04. (cited on pages vi, 5, 6, and 7)

[2] Texas Heart Institute. Heart anatomy, 2025. URL https://www.texasheart.org/heart-health/heart-information-center/topics/heart-anatomy/. Accessed: 2025-02-07. (cited on pages vi, 6, 7, and 9)

[3] Dakun Lai, Xinshu Zhang, Yuxiang Bu, Ye Su, and Chang-Sheng Ma. An automatic system for real-time identifying atrial fibrillation by using a lightweight convolutional neural network. *IEEE access*, 7: 130074–130084, 2019. (cited on pages vi, ix, 11, 17, 21, 79, 80, 87, 90, 91, 92, and 94)

[4] Putri Madona, Rahmat Ilias Basti, and Muhammad Mahrus Zain. Pqrst wave detection on ecg signals. *Gaceta sanitaria*, 35:S364–S369, 2021. (cited on pages vi and 14)

[5] Ridwan Alam, Aaron D Aguirre, and Collin M Stultz. Qtnet: Deep learning for estimating qt intervals using a single lead ecg. In *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 1–4. IEEE, 2023. (cited on pages vi, 17, 18, 21, 79, 81, 89, 91, 92, and 94)

[6] Adrian Luca, Sasan Yazdani, Alain Viso, Jean-Marc Vesin, Giulio

Conti, and Angelo Auricchio. Ecg-derived markers to identify patients prone to atrial fibrillation. In *2016 Computing in Cardiology Conference (CinC)*, pages 977–980. IEEE, 2016. (cited on pages vi, ix, 19, 20, 21, 85, 86, 88, 89, 90, 93, and 95)

[7] Francesco Renna, Jorge Oliveira, and Miguel T Coimbra. Deep convolutional neural networks for heart sound segmentation. *IEEE journal of biomedical and health informatics*, 23(6):2435–2445, 2019. (cited on pages vii, 34, and 35)

[8] BioSPPy Developers. BioSPPy Signals ECG Module Documentation. URL `https://biosppy.readthedocs.io/en/stable/biosppy.signals.html#module-biosppy.signals.ecg`. Accessed: 2026. (cited on pages vii, 38, and 39)

[9] Jasmin Bogatinovski, Dragi Kocev, and Aleksandra Rashkovska. Feature extraction for heartbeat classification in single-lead ecg. In *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pages 320–325. IEEE, 2019. (cited on pages ix, 76, 78, and 92)

[10] Bo Fang, Junxin Chen, Yu Liu, Wei Wang, Ke Wang, Amit Kumar Singh, and Zhihan Lv. Dual-channel neural network for atrial fibrillation detection from a single lead ecg wave. *IEEE journal of biomedical and health informatics*, 27(5):2296–2305, 2021. (cited on

pages ix, 79, 80, and 92)

[11] Yu Liu, Bo Fang, Yue Zhao, and Junxin Chen. Ensemble learning for atrial fibrillation screening from a single lead ecg wave of wearable devices. In *2021 IEEE 3rd International Conference on Frontiers Technology of Information and Computer (ICFTIC)*, pages 590–594. IEEE, 2021. (cited on pages ix, 79, 81, 82, and 93)

[12] Qiao Li, Kim Barnett, Joel Xue, Dave Albert, F Deepanshi, Ali Rad, Reza Sameni, and Gari Clifford. Fusion-based qt interval estimation for enhanced long qt detection on a south african population using mobile ecg. *Authorea Preprints*, 2025. (cited on pages ix, 84, 85, and 93)

[13] Freddie Bray, Mathieu Laversanne, Hyuna Sung, Jacques Ferlay, Rebecca L Siegel, Isabelle Soerjomataram, and Ahmedin Jemal. Global cancer statistics 2022: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 74(3):229–263, 2024. (cited on page 1)

[14] Federico Viganego, Robin Singh, and Michael G Fradley. Arrhythmias and other electrophysiology issues in cancer patients receiving chemotherapy or radiation. *Current cardiology reports*, 18(6):52, 2016. (cited on page 1)

[15] Michał Gawlik, Jakub Michal Zimodro, Aleksandra Gąsecka,

Krzysztof J Filipiak, and Sebastian Szmit. Cardiac arrhythmias in oncological patients—epidemiology, risk factors, and management within the context of the new esc 2022 guidelines. *Current Oncology Reports*, 25(10):1107–1115, 2023. (cited on page 1)

[16] Robert Monsour, Angie Seo, Nicholas Wilcox, and Michael G Fradley. Diagnosis and management of atrial arrhythmias in cancer patients. *Current Cardiology Reports*, 27(1):103, 2025. (cited on page 1)

[17] Sunil T Mathew, Jigar Patel, and Satheesh Joseph. Atrial fibrillation: mechanistic insights and treatment options. *European journal of internal medicine*, 20(7):672–681, 2009. (cited on page 1)

[18] H. Calkins and et al. 2017 hrs/ehra/ecas/aphrs/solaece expert consensus statement on catheter and surgical ablation of atrial fibrillation. *Heart Rhythm*, 14(10):e275–e444, 2017. (cited on page 1)

[19] Craig T January, L Samuel Wann, Joseph S Alpert, Hugh Calkins, Joaquin E Cigarroa, Joseph C Cleveland, Jamie B Conti, Patrick T Ellinor, Michael D Ezekowitz, Michael E Field, et al. 2014 aha/acc/hrs guideline for the management of patients with atrial fibrillation: a report of the american college of cardiology/american heart association task force on practice guidelines and the heart rhythm society. *Journal of the American College of Cardiology*, 64(21):e1–e76, 2014. (cited

on page 1)

[20] Gregory Y. H. Lip, Laurent Fauchier, Saul B. Freedman, Isabelle Van Gelder, Andrea Natale, Carola Gianni, Stanley Nattel, Tatjana Potpara, Michiel Rienstra, Hung-Fat Tse, and Deirdre A. Lane. Atrial fibrillation. *Nature Reviews Disease Primers*, 2:1–26, 2016. (cited on page 1)

[21] Suranga Seneviratne, Yining Hu, Tham Nguyen, Guohao Lan, Sara Khalifa, Kanchana Thilakarathna, Mahbub Hassan, and Aruna Seneviratne. A survey of wearable devices and challenges. *IEEE Communications Surveys & Tutorials*, 19(4):2573–2620, 2017. (cited on page 2)

[22] Giuseppe Boriani, Jacopo F Imberti, Riccardo Asteggiano, Pietro Ameri, Davide A Mei, Michał Farkowski, Julian Chun, Josè Luis Merino, Teresa Lopez-Fernandez, and Alexander R Lyon. Mobile/wearable digital devices for care of active cancer patients: a survey from the esc council of cardio-oncology. *European Heart Journal-Digital Health*, 6(2):162–169, 2025. (cited on page 2)

[23] Vishy Mahadevan. Anatomy of the heart. *Surgery (Oxford)*, 36(2): 43–47, 2018. (cited on page 5)

[24] Grace Casaclang-Verzosa, Bernard J Gersh, and Teresa SM Tsang. Structural and functional remodeling of the left atrium: clinical and

therapeutic implications for atrial fibrillation. *Journal of the American College of Cardiology*, 51(1):1–11, 2008. (cited on page 8)

[25] Amin Deyranlou, Josephine H Naish, Christopher A Miller, Alistair Revell, and Amir Keshmiri. Numerical study of atrial fibrillation effects on flow distribution in aortic circulation. *Annals of biomedical engineering*, 48(4):1291–1308, 2020. (cited on page 9)

[26] Luna Fabricius Ekenberg, Dan Eik Høfsten, Søren M Rasmussen, Jesper Mølgaard, Philip Hasbak, Helge BD Sørensen, Christian S Meyhoff, and Eske K Aasvang. Wireless single-lead versus standard 12-lead ecg, for st-segment deviation during adenosine cardiac stress scintigraphy. *Sensors*, 23(6):2962, 2023. (cited on page 10)

[27] T. Health. 12-lead ecg vs. single-lead ecg: The superior choice. `https://www.tricog.com/blog/12-lead-vs-single-lead-ecg`, 2023. Reports sensitivity 44% (single-lead) vs. 79% (12-lead) for ST-elevation. (cited on page 10)

[28] Abdilahi Mohamoud, Joseph Jensen, and Kevin G Buda. Consumer-grade wearable cardiac monitors: What they do well, and what needs work. *Cleveland Clinic Journal of Medicine*, 91(1):23–29, 2024. (cited on page 10)

[29] E. Vergara-Escudero, S. Vieau, and P. A. Iaizzo. Basic ecg theory,

12-lead recordings and their interpretation. In *Handbook of Cardiac Anatomy, Physiology, and Devices*, pages 317–332. Springer Nature Switzerland, Cham, 2024. (cited on page 11)

[30] Hugh Calkins, Gerhard Hindricks, Riccardo Cappato, Young-Hoon Kim, Eduardo B. Saad, Luis Aguinaga, J. G. Akar, Nitish Badhwar, Josep Brugada, and et al. 2017 hrs/ehra/ecas/aphrs/solaece expert consensus statement on catheter and surgical ablation of atrial fibrillation. *Heart Rhythm*, 14(10):e275–e444, 2017. (cited on pages 11 and 87)

[31] H. Calkins and et al. 2017 hrs/ehra/ecas/aphrs/solaece expert consensus statement on catheter and surgical ablation of atrial fibrillation. *Heart Rhythm*, 14(10):e275–e444, 2017. (cited on page 11)

[32] National Center for Biotechnology Information. Pr interval (concept id: C3152251). `https://www.ncbi.nlm.nih.gov/medgen/463601`, 2026. Accessed April 2026. (cited on page 12)

[33] Min Cheng, Xiangfeng Lu, Jianfeng Huang, Shu Zhang, and Dongfeng Gu. Electrocardiographic pr prolongation and atrial fibrillation risk: a meta-analysis of prospective cohort studies. *Journal of cardiovascular electrophysiology*, 26(1):36–41, 2015. (cited on page 13)

[34] J Olbertz, RM Lester, and M Combs. Establishing normal ranges

for ecg intervals in a normal healthy population. *Age*, 25(10734):27, 2015. (cited on page 13)

[35] Mala C Mandyam, Elsayed Z Soliman, Alvaro Alonso, Thomas A Dewland, Susan R Heckbert, Eric Vittinghoff, Steven R Cummings, Patrick T Ellinor, Bernard R Chaitman, Karen Stocke, et al. The qt interval and risk of incident atrial fibrillation. *Heart rhythm*, 10(10): 1562–1568, 2013. (cited on page 13)

[36] Dominique Makowski, Tam Pham, Zen J Lau, Jan C Brammer, François Lespinasse, Hung Pham, Christopher Schölzel, and SH Annabel Chen. Neurokit2: A python toolbox for neurophysiological signal processing. *Behavior research methods*, 53(4):1689–1696, 2021. (cited on pages 19, 29, 81, and 89)

[37] Python Software Foundation. Python language reference, version 3.11. `https://www.python.org`, 2026. Accessed: 2026-02-06. (cited on page 25)

[38] Dominique Makowski, Tam Pham, Zachary J. Lau, Jan C. Brammer, François Lespinasse, Hung Pham, Christopher Schölzel, and Sophia A. Chen. Neurokit2: A python toolbox for neurophysiological signal processing. `https://neuropsychology.github.io/NeuroKit/introduction.html`, 2021. Accessed: January 2026. (cited on page 29)

[39] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015. (cited on page 34)

[40] Daniel Proano-Guevara, André Lobo, Cristina Oliveira, Cátia Isabel Costa, Ricardo Fontes-Carvalho, Hugo Plácido da Silva, and Francesco Renna. Bidirectional fiducial matching of electrocardiography and phonocardiography for multimodal signal quality assessment. (cited on page 34)

[41] Alena Kalyakulina, Igor Yusipov, Viktor Moskalenko, Alexander Nikolskiy, Konstantin Kosonogov, Nikolai Zolotykh, and Mikhail Ivanchenko. Lobachevsky university electrocardiography database. *Type: Dataset. Available online: https://physionet.org/content/ludb/1.0. 0/(accessed on 10 July 2021)*, 2020. (cited on page 35)

[42] Yilin Pan, Kexin Peng, Yichuan Jiang, Ping Yang, Beibei Du, and Yuquan He. Electrocardiographic changes in qtc interval and other parameters associated with osimertinib therapy. *Frontiers in Oncology*, 15:1612758, 2025. (cited on page 63)

[43] Matthew J Page, Joanne E McKenzie, Patrick M Bossuyt, Isabelle

Boutron, Tammy C Hoffmann, Cynthia D Mulrow, Larissa Shamseer, Jennifer M Tetzlaff, Elie A Akl, Sue E Brennan, et al. The prisma 2020 statement: an updated guideline for reporting systematic reviews. *bmj*, 372, 2021. (cited on page 72)

[44] Ruhi Mahajan, Rishikesan Kamaleswaran, John Andrew Howe, and Oguz Akbilgic. Cardiac rhythm classification from a short single lead ecg recording via random forest. In *2017 Computing in Cardiology (CinC)*, pages 1–4. IEEE, 2017. (cited on pages 76, 77, and 91)

[45] Andrii Yavorskyi, Bohdan Tyshchenko, and Taras Panchenko. Efficient ecg analysis with high f 1 score and low computation complexity. In *2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, volume 1, pages 348–352. IEEE, 2021. (cited on pages 76, 77, and 92)

[46] Jiapu Pan and Willis J Tompkins. A real-time qrs detection algorithm. *IEEE transactions on biomedical engineering*, (3):230–236, 2007. (cited on pages 77 and 89)

[47] Mateusz Soliński, Anna Perka, Jacek Rosiński, Michał Łepek, and Joanna Rymko. Classification of atrial fibrillation in short-term ecg recordings using a machine learning approach and hybrid qrs detection. In *2017 Computing in Cardiology (CinC)*, pages 1–4. IEEE,

2017. (cited on pages 83 and 93)

[48] Charilaos A Zisou, Georgios K Apostolidis, and Leontios J Hadjileontiadis. Investigation of the evolution of wavelet higher-order dynamics in atrial fibrillation. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 363–366. IEEE, 2022. (cited on pages 85, 86, 87, and 93)

[49] Pil-Sung Yang, Soorack Ryu, Daehoon Kim, Eunsun Jang, Hee Tae Yu, Tae-Hoon Kim, Jinseub Hwang, Boyoung Joung, and Gregory YH Lip. Variations of prevalence and incidence of atrial fibrillation and oral anticoagulation rate according to different analysis approaches. *Scientific reports*, 8(1):6856, 2018. (cited on page 87)

[50] Thomas Hennig, Philipp Maass, Junichiro Hayano, and Stefan Heinrichs. Exponential distribution of long heart beat intervals during atrial fibrillation and their relevance for white noise behaviour in power spectrum. *Journal of biological physics*, 32(5):383–392, 2006. (cited on page 87)

[51] Dakun Lai, Xinshu Zhang, Yifei Zhang, and Md Belal Bin Heyat. Convolutional neural network based detection of atrial fibrillation combing rr intervals and f-wave frequency spectrum. In *2019 41st annual international conference of the IEEE engineering in medicine*

and biology society (EMBC), pages 4897–4900. IEEE, 2019. (cited on page 90)