Diogo Bernardo Jacinto Tecelão

Bachelor of Science in Biomedical Engineering

Prediction of postoperative atrial fibrillation using the electrocardiogram: A proof of concept

Dissertation submitted in partial fulfillment
of the requirements for the degree of

Master of Science in
Biomedical Engineering

Adviser: Dr Peter H Charlton, Research Associate,
King's College London

Co-adviser: Dr Pedro Vieira, Assistant Professor,
Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa

Examination Committee

Chairperson: Dr Célia Maria Reis Henriques, Assistant Professor at FCT-NOVA
Raporteur: Dr Carla Maria Quintão Pereira, Assistant Professor at FCT-NOVA
Member: Dr Pedro Manuel Cardoso Vieira, Assistant Professor at FCT-NOVA

FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

September, 2018

Prediction of postoperative atrial fibrillation using the electrocardiogram: A proof of concept

This document was created using the (pdf)LaTeX processor, based in the "novathesis" template[1], developed at the Dep. Informática of FCT-NOVA [2].
[1] https://github.com/joaomlourenco/novathesis [2] http://www.di.fct.unl.pt

Para a minha avó São e tio Melro.

Acknowledgements

I would like to acknowledge and thank the following important people who have supported me throughout this work.

Firstly, I would like to express my sincere gratitude to my supervisor Dr Peter Charlton. Peter has been an inspiration and a role model, having always helped, advised, and encouraged me unconditionally. The teachings and friendship Peter has provided me makes all this journey and thesis worth it. Secondly, I would like to thank Dr Pedro Vieira, whom I look up to, and whose support and advices were important.

I am also thankful to my close family who have always supported me. Without their encouragement and faith in me, I would never get this far. I would like to thank all my friends, who made all this journey worth and a lot more fun. A special thank you to: Filipe Valadas, who has greatly supported and inspired me (may the *patadas* and occlumency never end); André Pita, for all his friendship and prophylactic thoughts; and Ana Maria, for all her love and support.

Finally, I would like to thank Professor Rui Cardoso, my cousin Nuno, and Professor Antonieta Freire for all their help during my scholar journey. Without them I would not get this far.

Choose a job you love and you will never have to work a day in your life.

Abstract

Hospital patients recovering from major cardiac surgery are at high risk of postoperative atrial fibrillation (POAF), an arrhythmia which can be life-threatening. With the development of a tool to predict POAF early enough, the development of the arrhythmia could be potentially prevented using prophylactic treatments, thus reducing risks and hospital costs. To date, no reliable method suitable for autonomous clinical integration has been proposed yet.

This thesis presents a study on the prediction of POAF using the electrocardiogram. A novel P-wave quality assessment tool to automatically identify high-quality P-waves was designed, and its clinical utility was assessed. Prediction of paroxysmal atrial fibrillation (AF) was performed by implementing and improving a selection of previously proposed methods. This allowed to perform a systematic comparison of those methods, and to test if their combination improved prediction of AF. Finally, prediction of POAF was tested in a clinically relevant scenario. This included studying the 48 hours preceding POAF, and automatically excluding noise-corrupted P-waves using the quality assessment tool.

The P-wave quality assessment tool identified high-quality P-waves with high sensitivity (0.93) and good specificity (0.84). In addition, this tool improved the ability to predict AF, since it improved the precision of P-wave measurements. The best predictors of AF and POAF were measurements of the variability in P-wave time- and morphological features. Paroxysmal AF could be predicted with high specificity (0.93) and good sensitivity (0.82) when several predictors were combined. Furthermore, POAF could be predicted 48 hours before its onset with good sensitivity (0.74) and specificity (0.70). This leaves time for prophylactic treatments to be administered and possibly prevent POAF. Despite being promising, further work is required for these techniques to be useful in the clinical setting.

Keywords: Postoperative atrial fibrillation, electrocardiogram, P-wave, prediction.

Resumo

Pacientes a recuperar de cirurgia cardíaca estão em alto risco de desenvolver fibrilação auricular pós-operativa (FAPO), uma arritmia potencialmente fatal. O desenvolvimento de uma ferramenta para prever a FAPO permitiria evitar a arritmia através de tratamentos profiláticos, reduzindo assim os riscos e custos hospitalares associados. Até ao momento não existe nenhuma técnica fidedigna capaz de integração autónoma no meio clínico.

Esta tese apresenta um estudo sobre a previsão da FAPO através do eletrocardiograma. Foi criada uma ferramenta inovadora que permite a identificação automática de ondas-P de alta qualidade, e a sua utilidade foi testada no contexto clínico. A previsão da fibrilação auricular (FA) paroxística foi realizada através da implementação e melhoria de uma seleção de métodos propostos na literatura. Isto permitiu a realização de uma comparação sistemática entre esses métodos, e testar se a sua combinação melhora a previsão da FA. Finalmente, a previsão da FAPO foi realizada num cenário clinicamente relevante: estudando as 48 horas anteriores ao começo da arritmia, e excluindo automaticamente ondas-P corrompidas por artefactos através da ferramenta de avaliação de qualidade.

A ferramenta de avaliação de qualidade das ondas-P identificou ondas-P de alta qualidade com alta sensibilidade (0,93) e boa especificidade (0,84). Além disso, esta ferramenta melhorou a capacidade de prever a FA, através do aumento da precisão das medições retiradas das ondas-P. Os melhores previsores da AF e FAPO foram medidas de variabilidade dos intervalos e morfologia da onda-P. A FA paroxística foi prevista com alta especificidade (0,93) e boa sensibilidade (0,82) através da combinação de vários métodos. A FAPO foi prevista com 48 horas de antecedência com boa sensibilidade (0,74) e especificidade (0,70). Esta antecedência permite que os tratamentos profiláticos sejam administrados e possivelmente evitem a FAPO. Apesar dos resultados promissores, trabalho futuro é necessário para que estas técnicas sejam úteis na prática clínica.

Palavras-chave: Fibrilação auricular pós-operativa, electrocardiograma, onda P, previsão.

Publications

This thesis includes work that will be reported in different publications.

Currently, the work presented in Chapter 4 will be presented in the 5th International Electronic Conference on Sensors and Applications (ECSA-5). This will be published as a conference paper, entitled "Automated P-wave quality assessment for wearable sensors".

Contents

List of Figures

List of Tables

Acronyms

ABP	Arterial Blood Pressure.
ADI	Amplitude Dispersion Index.
AF	Atrial Fibrillation.
AFPDB	Atrial Fibrillation Prediction Database.
AUC	Area Under the Curve.
AV	Atrioventricular.
bpm	beats per minute.
CABG	Coronary Artery Bypass Grafting.
CCI	Correlation Coefficient Index.
CinC	Computers in Cardiology.
CTM	Central Tendency Measurement.
DTW	Dynamic Time Warping.
ECG	Electrocardiogram.
ESC	European Society of Cardiology.
FFT	Fast Fourier Transform.
HR	Heart Rate.
HRV	Heart Rate Variability.
ICU	Intensive Care Unit.
IV	Intravenous.
LE	Location Error.
LoA	Limits of Agreement.

MAE	Mean Absolute Error.
MIMIC-III	Medical Information Mart for Intensive Care III.
NPV	Negative Predictive Value.
PAF	Paroxysmal Atrial Fibrillation.
POAF	Postoperative Atrial Fibrillation.
PPG	Photoplethysmogram.
PPV	Positive Predictive Value.
PQI	P-wave Quality Index.
PT	Phasor Transform.
PV	Pulmonary Veins.
QTDB	QT Database.
RMS	Root Mean Square.
ROC	Receiver Operating Characteristic.
SA	Sinoatrial.
SD	Standard Deviation.
SNR	Signal-to-noise ratio.
SQI	Signal Quality Index.
WF	Warping Function.
WI	Warping Index.

Chapter 1

Introduction

Atrial fibrillation (AF) is the most common cardiac arrhythmia in clinical practice, and a well-recognized complication of cardiac surgery [1]. Patients recovering from major cardiac surgery are at high-risk of developing postoperative AF (POAF), with an incidence of 11% - 62%, depending on the type of surgery [2–7]. Importantly, the frequency of this arrhythmia appears to be increasing, most likely due to the increasing age of patients undergoing cardiac surgery [8].

Even though AF is self-limiting in nature, this arrhythmia is associated with increased risk of stroke and hemodynamic compromise, mortality, increased length of hospital stay, and health care costs [9–11]. In the long term, patients who develop POAF have a twofold increase in cardiovascular mortality [12], and a substantial increase in the risk of future AF and ischemic stroke, compared to patients who remain in sinus rhythm after surgery [13–15].

Prophylactic treatment has been shown to be effective in the prevention of POAF [16–20]. However, it is not routinely used because indiscriminate use of prophylactic agents does not outweigh the risks of side effects [21, 22] and has added costs [22]. Therefore, if a tool could be developed to predict the onset of POAF early enough, then the development of the arrhythmia could potentially be prevented using prophylactic treatments, thus reducing risks and hospital costs.

This issue has been approached both pre- and postoperatively, achieving encouraging yet insufficient results. The P-wave from the electrocardiogram (ECG) has been widely studied since it reflects atrial depolarization. Preoperative risk stratification has used AF risk factors, with or without ECG measurements, to identify patients at risk of developing AF [23]. Even though these studies have identified important factors that predispose individuals to the arrhythmia, their clinical application is limited by their moderate performance. More recently, long-term continuous monitoring has led to the identification

of subtle P-wave alterations that were found to predict AF at least 2 hours before its onset with high accuracy [24–27].

Although considerable progress has been achieved in the prediction of AF, no reliable method suitable for autonomous clinical integration has been proposed yet. Several additional steps are required to implement such systems in clinical practice. The first step concerns the improvement of predictive methods and evaluation in clinically relevant scenarios. This includes the prediction of POAF sufficiently long before its onset when prophylactic treatment is beneficial and the inclusion of heart rhythms other than sinus rhythm, which are commonly found in the postoperative setting. Secondly, such prediction techniques need to be adapted for unsupervised use during continuous monitoring, where ECG data is susceptible to artefacts that lead to inaccurate measurements.

1.1 Thesis Goals

This thesis aimed to address several of the steps required before techniques to predict POAF can be implemented in the clinical setting. The hypothesis of this project is that:

ECG signals can be analysed to predict the onset of POAF with sufficient warning time to allow prophylactic treatment to be administered.

Several goals were identified in order to achieve the overall aim as follows:

- **Design a P-wave quality assessment tool, and assess its performance and clinical utility**. This tool will automatically assess P-wave quality. Performance assessment will include the ability to distinguish between high and low quality data. The clinical utility of the tool will be assessed by testing if the developed tool improves the precision of several P-wave feature measurements and prediction of AF.

- **To identify, implement and improve previously proposed methods for predicting AF using the ECG signal, and assess their performance using a database containing records from healthy subjects and patients before the onset of AF**. Performance assessment will include the ability to distinguish between controls and subjects who subsequently develop AF, and to predict the imminent onset of the arrhythmia in subjects who develop AF.

- **Assess the performance of the previously implemented methods to predict POAF on a real world clinical database**. Prediction of POAF will be performed in a clinically relevant scenario: 48 hours before the arrhythmia and without supervision.

The novelties of this study are that it is the first to predict POAF during a clinically realistic scenario, and that it is the first investigating the 48 hours prior to the onset of the arrhythmia. Furthermore, this study is the first to perform a systematic comparison of previously proposed methods, and test if they can be used in combination to improve

prediction of AF. Finally, a novel P-wave quality assessment tool is presented, which allows the AF prediction to be conducted without supervision.

1.2 Thesis Outline

The thesis is structured as follows. Chapter 2 presents the clinical background, including the cardiovascular system, the ECG signal, and an overview of both AF and POAF. In Chapter 3 the state of the art on the prediction of AF is presented, including preoperative AF risk stratification and postoperative prediction of POAF. In chapter 4 a novel P-wave quality assessment tool is presented, and its performance and clinical utility are assessed. Chapter 5 presents a study on the ability of several proposed methods for predicting paroxysmal AF. Chapter 6 presents a study of the performance of these methods when predicting POAF in a clinically relevant scenario. Finally, Chapter 7 presents a summary of the achievements of this thesis and directions for future work.

Chapter 2

Clinical Background

In this chapter the clinical background to this thesis is presented. Firstly, the cardiovascular system is presented, including a physiological monitoring technique: the electrocardiogram (ECG). This is followed by an overview of atrial fibrillation (AF) and postoperative atrial fibrillation (POAF), including their epidemiology, mechanisms, risk factors and the current treatment methodologies.

2.1 The cardiovascular system

The heart consists of four muscular chambers. The left and right ventricles are the main pumping chambers, whilst the less muscular left and right atria deliver blood to their respective ventricles. Four major valves in the healthy heart direct blood flow in a forward direction and prevent backward leakage. The atrioventricular valves (tricuspid and mitral) separate the atria and ventricles, whereas the semilunar valves (pulmonary and aortic) separate the ventricles from the great arteries [28]. From a functional point of view, the deoxygenated blood is delivered to the heart through the inferior and superior venae cavae, and enters the right atrium. Flow continues through the tricuspid valve into the right ventricle. Contraction of the right ventricle pumps the blood across the pulmonary valve to the pulmonary artery and lungs, where carbon dioxide is released and oxygen is absorbed. The oxygen-rich arterial blood returns to the heart through the pulmonary veins to the left atrium and then passes through the mitral valve into the left ventricle. Contraction of the left ventricle propels the arterial blood through the aortic valve into the aorta, from where it is distributed to all other tissues across the body [28].

Cardiac contraction relies on the organized flow of electrical impulses through the heart's impulse-conducting system. The normal heartbeat begins at the sinoatrial (SA) node, located at the junction of the right atrium and the superior vena cava (Figure

2.1). The wave of depolarization spreads rapidly through the right and left atria and then reaches the atrioventricular (AV) node, where it encounters an expected delay (of approximately 0.1 seconds). This pause in conduction is beneficial in the sense that allows the atria time to contract and fully empty their contents and serves as a "gatekeeper" of conduction from atria to ventricles, limiting the rate of ventricular stimulation during atrial arrhythmias [28]. After traversing the AV node, the impulse then travels rapidly through the bundle of His and into the right and left bundle branches. Those, in turn, divide into the Purkinje fibres, which radiate toward the myocardial fibres, stimulating them to depolarize and contract, pumping the blood through the respective valves [28].

2.1.1 The electrocardiogram

The aforementioned electrical activity of the heart results in small currents within the body and potential differences on the surface of the skin, which can be measured using skin electrodes. Each healthy heartbeat is represented on the ECG by three major deflections that record the sequence of electrical propagation. The P-wave represents the atrial depolarization, and is followed by a period of baseline resulting from the conduction delay at the AV node (PR interval). The second and more pronounced deflection on the ECG is the QRS complex, which represents the ventricular depolarization. After a brief delay during baseline, the T-wave surges representing ventricular repolarization (Figure 2.1). Occasionally, an additional small deflection named the U-wave follows, representing late phases of ventricular repolarization [28].

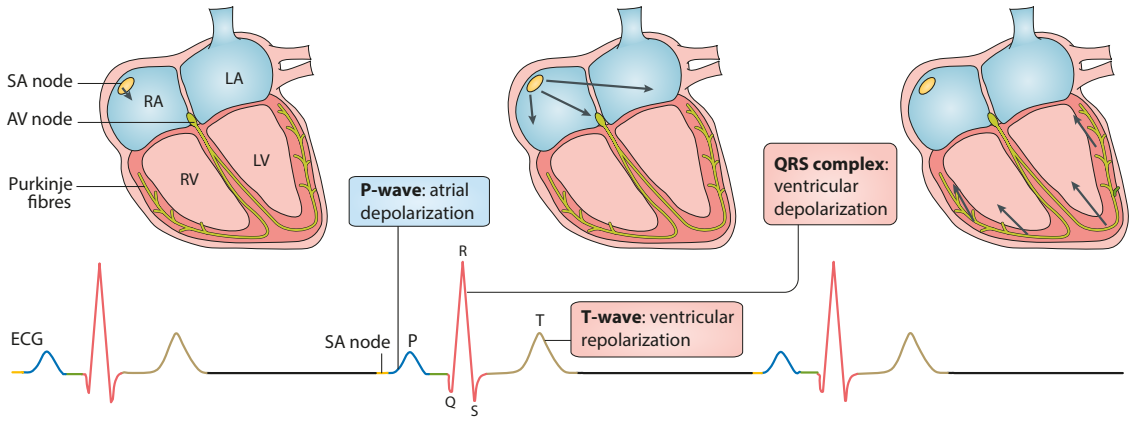

Figure 2.1: Representation of the electrical activity of the heart during sinus rhythm (top), and corresponding electrocardiogram (ECG; bottom). Adapted from Lip *et al.* [11].

2.2 Atrial Fibrillation

AF is a supraventricular arrhythmia with an atrial discharge rate so fast (350 to 600 discharges per minute) that distinct P waves are not discernible on the ECG [28], resulting in low-amplitude baseline oscillations [29]. As many of the atrial impulses encounter

refractory tissue at the AV node, allowing only some of the depolarization to be conducted to the ventricles in a very irregular fashion, the result is an "irregularly irregular" rhythm (Figure 2.2). The average ventricular rate in untreated AF is approximately 140 to 160 beats per minute (bpm) [28].

This disorder is the most common cardiac rhythm disorder, with an increasing global prevalence and incidence [1]. The most serious complications of AF include stroke, systemic thromboembolism, heart failure, and dementia, contributing to increased mortality and morbidity, and entailing high costs for families and society. The current economic burden of AF in Europe is substantial, with widespread implications for the planning of national health care systems [30], as the rate of hospitalization and treatment costs are increasing in epidemic proportions [31].

2.2.1 Epidemiology

AF is the most common arrhythmia found in clinical practice, affecting approximately 4.5 million in the European Union and 2.2 million in North America. Furthermore, AF is responsible for nearly 33% of arrhythmia-related hospitalizations [28]. There are almost five million new cases of AF every year, with the number of affected individuals expected to increase continuously in an exponential fashion [11, 32, 33]. The cause of this global increase might be due to an ageing population, since age is the strongest risk factor for developing AF. Furthermore, improvements in how conditions associated with AF are managed might also contribute to the increased AF rates, as survival of these conditions can be associated with cardiac damage, which also predisposes to AF [11].

AF is linked to numerous serious complications: a four- to five-fold increased risk of stroke [11, 34, 35], a three-fold increased risk of heart failure, and a nearly two-fold increased risk of all-cause mortality [36]. Strokes related to AF are associated with greater mortality, disability, with longer hospital stays, and lower rates of discharge to a patient's own home, compared with strokes that are not associated with AF [37]. Patients with AF generally experience increased morbidity and more admissions to hospital than those who do not have the arrhythmia. Increasing evidence also suggests an association between AF and an increased risk of 'premature' dementia [11].

2.2.2 Mechanisms

The central feature of AF is very rapid and irregular atrial activity. This generally requires a trigger to be initiated, which is typically a focal spontaneous firing. These triggers most commonly arise as rapid firing foci in sleeves of atrial muscle that extend into the pulmonary veins (PV), although it can also emerge from non-PV foci [11]. In paroxysmal AF (i.e., sudden, unpredictable episodes), AF is often initiated by these rapid discharges from the PV, reflected in premature atrial complexes [38–41]. The mechanism maintaining the arrhythmia often arises in what is commonly referred to as vulnerable substrate. This substrate might be associated with the several risk factors, such as genetic predisposition,

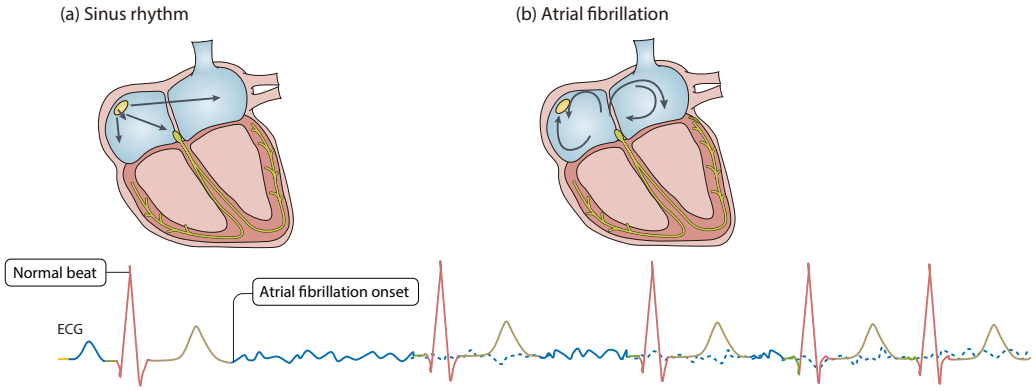

Figure 2.2: Electrical conduction during sinus rhythm and atrial fibrillation. Schematic diagrams of cardiac mechanisms (top of each panel) and electrocardiograms (ECGs; bottom of each panel) in (a) normal sinus rhythm and (b) atrial fibrillation. Adapted from Lip *et al.* [11].

cardiac remodelling caused by heart disease, and/or altered regulation by neurohormonal factors such as autonomic imbalance and overactive thyroid function.

Each of the following three principal mechanisms can maintain the chaotic rhythm that comprises AF [11] (Figure 2.3):

1. One or more rapidly firing atrial ectopic foci may be present, with irregular conduction towards the rest of the atria, producing irregular fibrillary-like activity.

2. One or a small number of primary re-entry circuits (or rotors) may produce rapid local activation, with fibrillary conduction causing AF.

3. Multiple functional re-entry waves with irregular patterns and no consistent activation pattern, maintaining the disordered AF rhythm.

In fact, slowed conduction velocity together with decreased cell refractory periods are believed to provoke and maintain AF, as such abnormalities result in a non-uniform and anisotropic atrial conduction which plays a major role in the initiation of re-entry.

The electromechanical consequences of AF have serious clinical implications. The absence of effective atrial contraction increases the risk of blood coagulation and thrombosis. Furthermore, the rapid and irregular ventricular rate generated by the arrhythmia reduces the efficiency of ventricular contraction, affecting hemodynamics and, ultimately, causing heart failure [11, 42].

2.2.3 Risk factors

Most risk factors are consistently associated with an increased risk of developing AF across all ethnic groups [11], even though ethnicity itself is reported to affect the incidence of the arrhythmia [43, 44]. AF is more common in elderly male individuals with cardiovascular abnormalities, such as hypertension, ischemic heart disease, or forms of

structural heart disease. Obesity and obstructive sleep apnea, although related, have been found to independently increase the risk for AF. The most common temporary causes are excessive alcohol consumption, open heart or thoracic surgery (POAF), myocardial infarction, pericarditis, myocarditis, and pulmonary embolism. The most common correctable cause if hyperthyroidism [28].

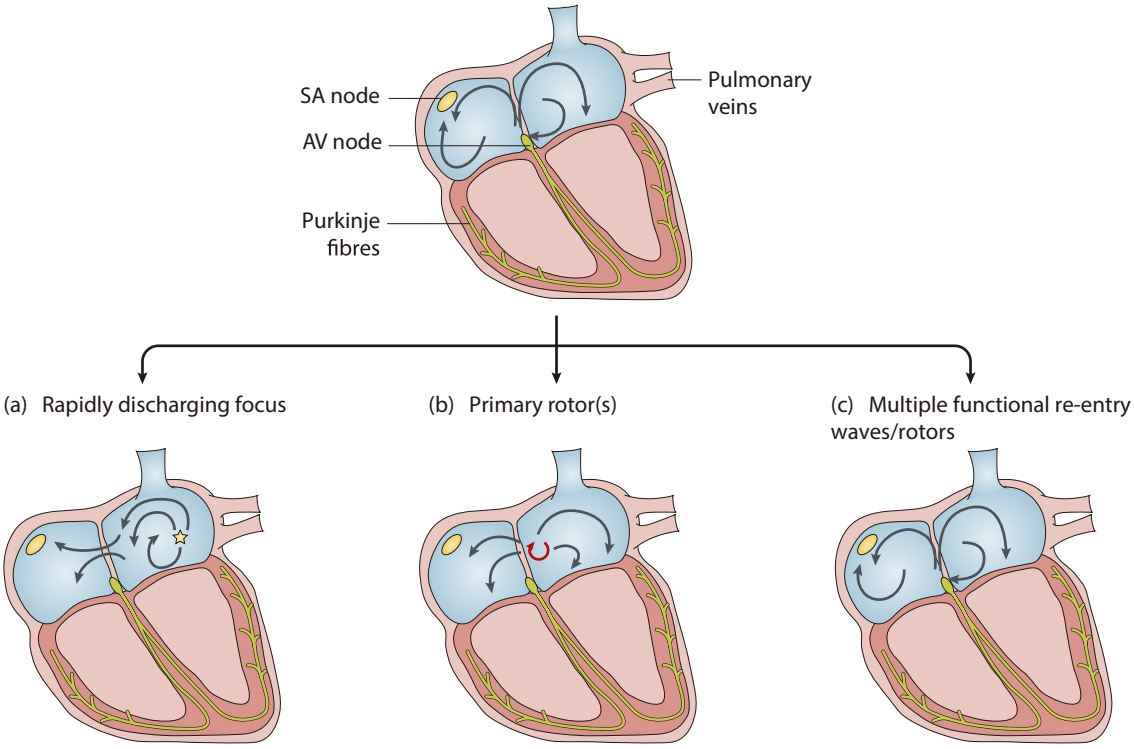

Figure 2.3: Mechanisms responsible for atrial fibrillation maintenance. Ectopic electrical impulses that propagate throughout the atrial myocardium in a disordered way can be maintained through a variety of mechanisms: (a) a rapidly discharging atrial focus, (b) primary re-entrant rotor(s), (c) multiple functional re-entry circuits. AF, atrial fibrillation; AV, atrioventricular; LA, left atrium; LV, left ventricle; PV, pulmonary vein; RA, right atrium; SA, sinoatrial. Adapted from Lip *et al* [11].

2.3 Postoperative Atrial Fibrillation

POAF is the most frequent complication of cardiac surgery [9], with a prevalence dependent on the type and technique of surgery, patient characteristics, method of arrhythmia surveillance, and definition of the arrhythmia [20]. Importantly, the incidence of POAF is increasing, most likely due to the increasing number of elderly subjected to cardiac surgery. Similarly to other types of AF, it affects patient well-being, by increasing the risks of stroke, hemodynamic compromise, and mortality [13, 14]. Furthermore, it increases the hospital length of stay, and incurs additional treatment costs [9].

2.3.1 Epidemiology

The incidence of POAF ranges from 11% to 62%, with the highest rate after combined coronary artery bypass grafting (CABG) and valve surgery [7, 45]. POAF may also occur in up to 40% in patients undergoing CABG surgery [4–6], 35% to 40% after valvular surgery [4, 5, 46], and 11% to 24% after cardiac transplantation [4, 5]. Initial onset of AF occurs most commonly on postoperative days two or three [6], and the majority of episodes occur within the first six days following cardiac surgery [10]. The incidence in patients undergoing non-cardiothoracic surgery varies between 0.3% and 30%, depending greatly on the specific surgical procedure [47–50].

POAF is associated with increased incidence of other postoperative complications, stroke being the most recognized, with a threefold risk increase [4, 51]. There is evidence of association between POAF and myocardial infarction, congestive heart failure, ventricular arrhythmias, renal insufficiency, hypotension, pulmonary edema, pneumonia, mediastinitis or deep sternal wound infection, sepsis, and harvest site infection. (It is important to note that these associations do not necessarily indicate a causal relationship between POAF and the mentioned complications, as it may just be an epiphenomenon [8]). Patients who develop POAF are also associated with the need for a permanent pacemaker, inotropic medications, prolonged ventilation, readmission to the ICU, and consequent increased hospital length of stay [51]. The impact of POAF on hospital resources is significant, with an extra cost of care of approximately $10,000 (8,700€) per patient in the United States [3]. In the long term, patients with an episode of POAF have a twofold increase in cardiovascular mortality [12], and a substantial increase in the risk of future AF and ischemic stroke, compared to patients who remain in sinus rhythm after surgery [13–15].

2.3.2 Mechanisms

The electrophysiological mechanisms underlying AF after cardiac surgery are not fully understood. However, it is believed that patients who develop POAF may already have preexisting atrial substrate for this arrhythmia before surgery [8], such as atrial fibrosis or dilation, with surgery being the trigger for AF. In patients undergoing cardiothoracic surgery this trigger can be related to the intraoperative trauma, manipulation of the heart, local inflammation (with or without pericarditis), elevated atrial pressure due to postoperative ventricular stunning, and the rapid return of atrial temperature after cardioplegia [45, 52]. In all-type surgical patients, AF can also be related to specific factors, such as direct cardiac stimulation from the perioperative use of catecholamines or reflex sympathetic activity from volume loss, anemia, fever, hypo- or hyperglycemia, and electrolyte disturbances [45]. There is also evidence that inflammation, both systemic and local, may play a role in the pathogenesis of POAF [52]. These surgical factors may therefore trigger POAF in susceptible patients through dispersion of atrial refractoriness [53], non-uniform atrial conduction [54], or increased premature atrial complexes [39]

(Figure 2.4).

2.3.3 Risk factors

Documented risk factors for the development of POAF can be divided into preoperative, perioperative and postoperative. In addition to common AF risk factors such as increased age and cardiovascular disease, other preoperative factors such as prior history of AF and hypothyroidism have been associated with the postoperative arrhythmia. Intraoperative variables include type of surgery, aortic cross-clamp time, the early return of atrial electrical activity after cardioplegia, atrial ischemia, systemic and pericardial inflammation, systemic hypothermia, and electrolytic imbalances, such as hypokalemia and hypomagnesemia. Postoperative conditions associated with POAF include respiratory compromise, which can be due to postoperative pneumonia, and hypotension.

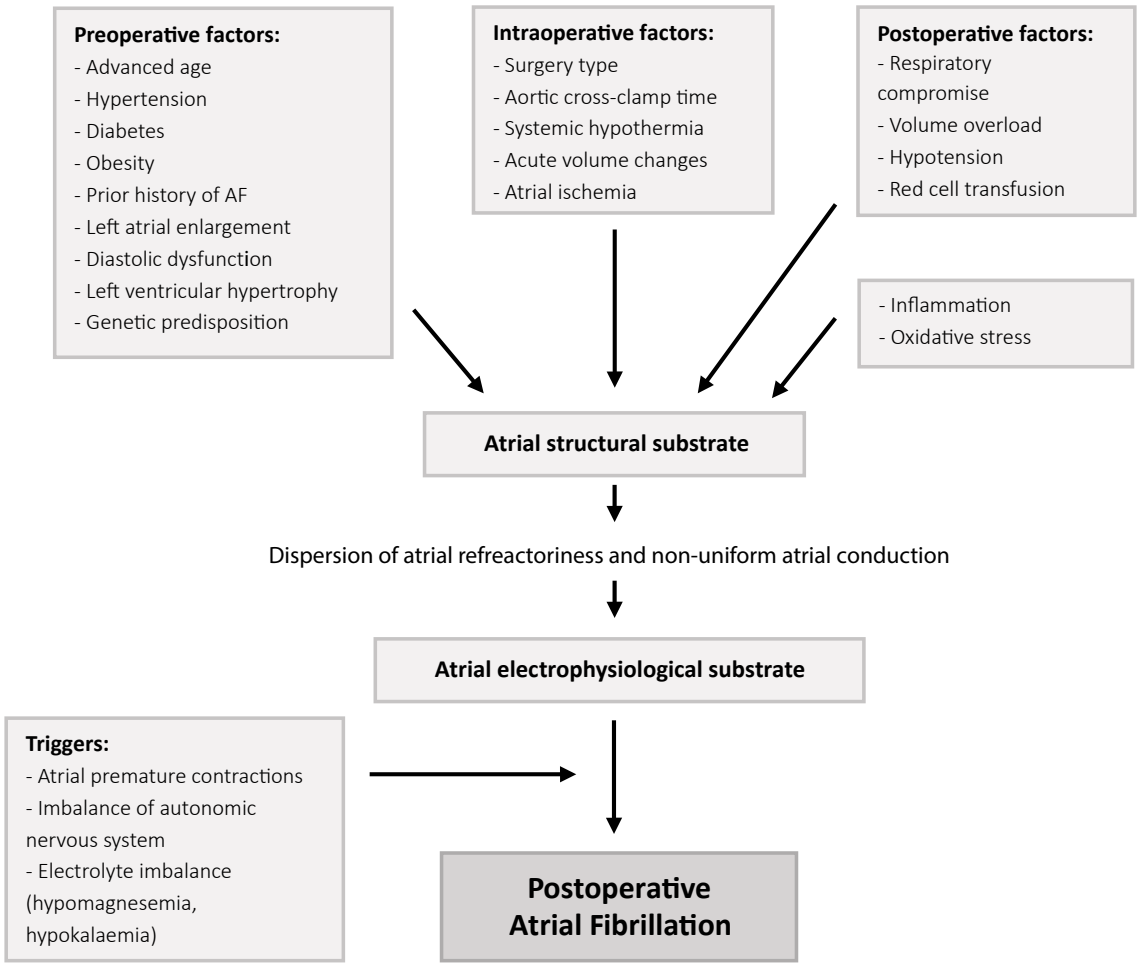

Figure 2.4: Pathogenesis of postoperative atrial fibrillation. It is believed that patients who develop POAF may already have preexisting atrial substrate that, together with intra- and post-operative factors, may trigger POAF through dispersion of atrial refractoriness, non-uniform atrial conduction, or increased premature complexes. Adapted from Echahidi *et al* [9].

2.3.4 Prevention

Many studies have evaluated the effectiveness of pharmacologic and non-pharmacologic interventions to prevent the onset of POAF, and therefore decrease its growing incidence. Current practice is based on the 2016 European Society of Cardiology (ESC) guidelines [55], and is presented bellow.

Beta-blockers. Beta-Adrenergic blockade helps reduce the effect of increased sympathetic activation, which is believed to be a significant contributor to the development of AF in the postoperative period, and to be present in patients undergoing cardiac surgery [9]. The preoperative initiation of beta-blockers has been demonstrated to be more effective than postoperative initiation. These drugs have the additional benefit of controlling the ventricular rate in the event of AF. Due to the extensive evidence of benefits from beta-blockers, the 2016 ESC guidelines presented a class I recommendation to perioperative beta-blocker therapy in patients subjected to cardiac surgery [55]. Although it is cost-effective for the reduction of AF, no significant decreases in hospital length of stay [17, 56] and overall cost of care have been observed [56]. The anti-arrhythmic effect of beta-blockers peaks immediately when administered intravenously (IV), and after 1 to 1.5 hours when administered orally [57]. Adverse effects include bradycardia, fatigue, bronchospasm, hypotension, and aggravation of heart failure [58, 59].

Amiodarone. Amiodarone has alpha- and beta-adrenergic blocking properties that might attenuate the elevated sympathetic stimulation seen in patients undergoing cardiac surgery [9, 60]. This anti-arrhythmic drug was found to reduce the incidence of POAF compared to beta-blocker therapy in several meta-analyses, also reducing hospital length of stay [55]. Perioperative amiodarone should be considered as prophylactic therapy to prevent POAF (class IIa in 2016 ESC guidelines), particularly so in high-risk patients [61]. The anti-arrhythmic effect of amiodarone peaks between 1 minute and 1.5 hours after IV administration [62]. When administered orally, the drug concentration peaks after 3 to 7 hours [63, 64], but the anti-arrhythmic effects only start after 2/3 days to 1/3 weeks [62, 65] (the oral absorption rate is slow and variable because amiodarone is highly lipophylic [58]). Adverse effects include symptomatic bradycardia and heart block in patients with pre-existing sinus or AV node disease, dose-related pulmonary toxicity, and hypo- or hyperthyroidism [59].

Although not present in the 2016 ESC guidelines, other therapies have also been studied, showing a decrease in POAF. Pharmacologic therapies include magnesium, colchicine, and corticosteroids [55]. A non-pharmacologic approach includes prophylactic atrial pacing, which acts through the suppression of atrial premature complexes and dispersion of atrial refractoriness, which are known to facilitate the onset of POAF [41, 45]. Despite suggestions of its prophylactic effects, biatrial pacing has not gained widespread use [55].

2.3.5 Treatment

Although POAF can be transient and self-limiting, treatment is indicated for patients who remain symptomatic, are hemodynamically unstable, and develop cardiac ischemia or heart failure [9]. Conventional treatment strategies include restoration/maintenance of sinus rhythm, control of ventricular rate response, and prevention of thromboembolic events. The treatment of POAF is manly based on studies of patients undergoing cardiac surgery, with much lower evidence on non-cardiac surgery [55].

Rate control. The postoperative period is characterized by increased adrenergic stress, which might difficult ventricular rate control in patients with POAF. Short-acting beta-blockers are the most widely used therapy, except for patients with contraindications to such agents. For those, other AV nodal blocking agents such as the nondihydropyridine calcium-channel blocker, can be used. Amiodarone is also effective in controlling heart rate, being further associated with improved hemodynamic status [9].

Rhythm control. In symptomatic patients or in cases of difficult control of ventricular response, cardioversion (i.e., conversion to sinus rhythm) is preferred. Chemical cardioversion with agents such as amiodarone, procainamide, ibutilide, and sotalol might be effective. Electrical cardioversion should be urgently performed when POAF results in hemodynamic instability, acute heart failure, or myocardial ischemia. This shock-procedure is electively used to immediately restore sinus rhythm after the first onset of AF when a pharmacologic attempt has failed [9]. In postsurgical patients, electrical cardioversion may be done either externally (transthoracic) or internally, using low-energy transvenous electrodes or epicardial wires placed during surgery. As atrial stunning persists after cardioversion, anticoagulation is recommended for the three to four weeks after conversion procedure. Thromboembolism is a major concern of cardioversion, either pharmacologically or electrically, particularly when POAF has been present for more than 48h [9].

Catheter ablation is an alternative approach. It is a well-established and commonly used therapeutic option for managing patients with symptomatic AF. It consists of ablating the site of anatomical reentry or fibrillary activity via a catheter that applies radiofrequency current to heat and destroy the tissue. However, this procedure is not recommended for management of POAF patients after thoracic surgery [66], since: 1) several factors that promote POAF cannot be addressed simply by this technique [29]; 2) it is common for AF to recur in the two- to three-month post-ablation healing phase, making it an inappropriate strategy for the control of acute, symptomatic AF, such as that which occurs in the postoperative setting; and 4) its efficacy is modest, and it is associated with risk of complications [66].

Thromboembolism prevention. Thromboembolic events are one of the most hazardous complications of AF, which can be reduced with anticoagulation. Moreover, oral anticoagulants at discharge have been associated with a reduction in long-term mortality in POAF. However, anticoagulation is associated with increased risk of bleeding or cardiac

tamponade when administered in the postoperative period [67]. Therefore, the indication and timing of this therapy in POAF patients should take into consideration the risk of postoperative bleeding. Aspirin is a suitable substitute for patients at low risk of thromboembolic events [8]. It is important to note that even though stroke prevention is central to AF management, data on anticoagulation prophylaxis in the postoperative setting are significantly lacking, and management is often based on evidence from nonsurgical data [45].

Chapter 3

Prediction of Postoperative Atrial Fibrillation: State of the Art

This chapter presents the state of the art on the prediction of postoperative atrial fibrillation (POAF). It is divided into two parts: preoperative atrial fibrillation (AF) risk stratification and postoperative prediction of AF. The first describes risk stratification efforts to predict POAF using either risk factors, preoperative electrocardiogram (ECG) features, or both combined. The second contains recent methodologies for analysing the ECG on a beat-to-beat basis to predict the onset of atrial fibrillation.

3.1 Preoperative AF risk stratification

An important aspect of POAF management is the identification of risk factors, and development of accurate prediction models. Risk stratification helps identifying patients who are most likely to develop POAF, and therefore most likely to benefit from targeted prophylactic treatment, thereby improving patient outcomes and reducing healthcare costs.

Several risk scores have been proposed for predicting the development of POAF, based on common risk factors that are significantly associated with the occurrence of the arrhythmia in their respective derivation cohorts (Table 3.1). Even though the predictive value of some of the models is reasonably good, they cannot be considered strong models, limiting their clinical application in risk stratification. Moreover, as in AF unrelated to surgery [43, 44, 68, 69], increased age is the only characteristic that has been consistently linked to an increased risk of POAF across all studies [6, 70–73]. The operation type has also been found to be a risk factor of POAF, although this has not been tested in some models [70, 71]. Importantly, no risk stratification has been performed in non-cardiac surgeries, further limiting the generalisability of the models to all surgery types.

Because structural factors, such as fibrosis, scarring, and dilatation of the atria, are predispositions to POAF, it has been hypothesized that electrophysiological measurements before surgery may help predict the development of the arrhythmia. Since then, over 25 ECG features extracted from the either P-wave or heart rate variability (HRV) have been proposed (Table 3.2). As the P-wave reflects atrial depolarization, the site responsible for the arrhythmia, it has been the most extensively studied component.

Table 3.1: Risk models for the prediction of incident atrial fibrillation after cardiac surgery.

Study	**Model inputs**	**Notes**	**Predictive value: c-statistic**
Mathew *et al.* [6]	• Age • Medical history of AF or COPD • Valve surgery • Withdrawal of beta-blockers or ACE inhibitors • Beta-blockers or ACE inhibitor treatment (pre- and/or postoperatively) • Postoperative potassium supplementation or nonsteroidal anti-inflammatory treatment	• The study also modeled a risk score for recurrence of AF and for the complications of POAF.	0.77
Helgadottir *et al.* [72]	• Age • Operation type • Standard EuroSCORE	• Patients with preoperative history of AF were excluded • No cross-validation was performed.	0.74
Mariscalco *et al.* [73]	• Age • Emergency operation • Preoperative aortic balloon • Left ventricular ejection fraction $< 30\%$ • Glomerular filtration rate $<$ 15mL/min or dialysis • Heart valve surgery • COPD	• The modeled POAF score was also predictive of hospital mortality.	0.65
Chua *et al.* [71]	• CHADS2: congestive heart failure, hypertension, age $\geqslant$ 75 years, type 2 diabetes, and previous stroke or transient ischemic attach (doubled). • CHA2DS2-VASc: congestive heart failure, hypertension, age $\geqslant$ 75 years (doubled), type 2 diabetes, previous stroke, transient ischemic attack or thromboembolism (doubled), vascular disease, age 65-75 years, and gender	• These scores were previously recommended and studied to guide antithrombotic therapy in patients with AF or atrial flutter. Each of the factors that can lead to atrial enlargement. • No validation using the area under the receiver-operating curve was performed.	n\a

AF, atrial fibrillation; ACE, angiotensin-converting enzyme; CI, confidence interval; COPD, congestive obstructive pulmonary disease; ECG, electrocardiogram; ms, milliseconds; n/a, not available; POAF, postoperative atrial fibrillation.

Even though some of the proposed methods have reached relatively good predictive values in terms of accuracy (the root mean square of the last 20ms of the signal-averaged P-wave combined with the signal-averaged P-wave duration provided the best result with an accuracy of 83% [74]), they are not sufficiently reliable for clinical use. Furthermore, due to the lack of a standard definition of P-wave onset and offset, P-wave time domain methods cannot be easily compared [23]. Nevertheless, the results obtained using these time domain indices are of clinical interest. Frequency domain indices and heart rate variability metrics have not shown to be useful *preoperative* predictors of POAF.

Table 3.2: Preoperative electrocardiogram-based methods for prediction of postoperative atrial fibrillation. These include P-wave and heart rate variability features in time, frequency and time-frequency (wavelet) domain.

ECG-component	Feature domain	Feature
P-wave	**Time**	Signal averaged duration [75–77]
		Duration [78–85]
		Isoelectric interval [76, 80]
		Signal-averaged root mean square [74]
		Root mean square [84]
		Terminal force [76]
		Spatial velocity [76]
		Dispersion [78, 79, 81–83]
		Amplitude [81, 82, 85]
		Variance [78]
		Area [85]
		P-axis [81]
		PR interval [81, 85]
	Frequency	Signal-averaged power [84, 86]
	Time-frequency	Mean and maximum energy [87]
Heart Rate Variability	**Time**	Mean [88]
		SDRR [88]
		rMSSD [88]
		NN50 [88]
		pNN50 [88]
		TINN [89]
		Triangular index [88]
		Embedded spectral entropy [88]
	Frequency	LF/HF [88]
	Geometric	Poincaré dispersion (SD1, SD2) [88]

3.2 Prediction of atrial fibrillation

Efforts to predict the imminent onset of AF started with the Computers in Cardiology 2001 challenge [90], which had the objective of developing a fully automated method to classify ECG signals as corresponding to just before the onset of paroxysmal AF or not. The developed methods mostly focused on the detection of atrial premature complexes using the inter-beat intervals [91–95] and heart rate variability metrics [91, 96–99], with a only single approach considering the QRS-complex morphology [96]. Even with a small dataset, results were insufficient and lacked accuracy (the best study achieved an accuracy of 80%).

Most recently, studies have investigated beat-to-beat analyses of the ECG during continuous monitoring, thus enabling real-time clinical prediction of AF. These methods can be divided into three stages, as depicted in Figure 3.1 and illustrated in Figure 3.2.

Figure 3.1: The three stages of atrial fibrillation (AF) onset prediction techniques, that perform beat-to-beat analyses of the ECG. Even though these methods have focused on the electrocardiogram (ECG) signal, they could also be applied to other signals.

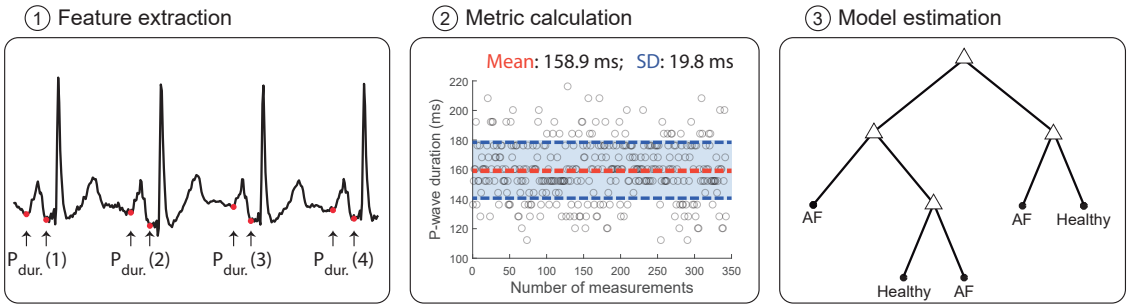

Figure 3.2: Illustration of the three stages of atrial fibrillation (AF) prediction algorithms. These techniques start by delineating the electrocardiogram signal and extract features (e.g. P-wave duration, as depicted). Then, those features are summarised into metrics (e.g. mean and standard deviation, as depicted, or more complex variability measurements), that are used to estimate models.

3.2.1 Feature extraction

The feature extraction stage derives a set of features from each beat of the ECG signal. Most studies have focused on the P-wave's intervals and morphology.

This stage relies on the identification of the ECG fiducial points. The initial step, common to all techniques, is the identification of the R peak of the QRS complex [100, 101]. This allows the calculation of several heart rate variability (HRV) metrics. The delineation of the P-wave was performed either using the dyadic wavelet [102, 103] or

phasor transform decomposition [104], with both methods showing high delineation accuracy. The use of such automated delineator allows not only the characterization of the P-wave, but also mitigates the lack of standard definition of the P-wave onset and end. Finally, several features are extracted using the P-wave fiducial points and signal.

3.2.2 Metric calculation

The metric calculation step, common to all methods, extracts several indices from each feature's time-series over a determined period of time (Table 3.3). This summarises each time-series of feature measurements as a single metric. Extracted metrics ranged from simple statistical metrics such as the mean, to more complex variability measurements.

3.2.3 Model estimation

This final step makes use of the previously calculated metrics to create a model to assess the likelihood of developing AF. Studies have relied on machine learning approaches, using the statistically significant metrics as inputs to the algorithms (Table 3.3).

Even though these studies have shown interesting and positive results, they rely on relatively small populations. Furthermore, the majority of the studies have based their models on signals from a time-frame too close to the onset of AF, which would not provide sufficient warning time for prophylactic treatment to be administered and take effect. Nonetheless, these studies have achieved encouraging results.

Overall, the beat-to-beat analyses have studied the variability of the extracted ECG features, and have obtained high prediction performance (with accuracy ranging from 70% and 100%). These studies have found that patients who will develop AF have greater variability in P-wave morphology and time-intervals (Figure 3.3) [25–27, 105, 107, 108], and that this variability increases as the onset approximates [25, 26, 105, 107, 108]. These subtle P-wave alterations are according to the atrial electrophysiological changes that precede AF, and suggest intermittently disturbed atrial conduction in patients close to start of the arrhythmia [25].

3.3 Conclusion

Prediction of POAF is preferable in the preoperative period because it allows the administration of drugs earlier. Indeed, the 2016 ESC guidelines recommend the administration of beta-blockers or amiodarone in the perioperative period [55]. However, results in preoperative risk stratification are limited. In contrast, recent studies have indicated that the prediction of POAF in the postoperative period is more promising, as these researches have obtained greater prediction performances (Table 3.3).

Even though considerable progresses have been achieved in the postoperative prediction of AF, no technique reliable for autonomous clinical integration has been proposed yet. Further work is required to improve the prediction of POAF from continuous and

Table 3.3: Studies that have performed beat-to-beat analyses of the electrocardiogram for predicting several types of atrial fibrillation, which would be suitable for use in continuous postoperative monitoring.

Study	**Cohort dimension**	**Time before AF onset**	**ECG features**	**Used metric**	**Model (accuracy)**
Martinez *et al.* [25]	24 PAF + 28 healthy	2h	• P-wave intervals • Heart rate	Linear regression slope of the variability	Linear discriminant (90.8 %)
Martinez *et al.* [26]	46 PAF + 53 healthy	2h	• P-wave morphology: area, conduction velocity, arc-length	Linear regression slope of the variability	Decision tree (86 3%)
Martinez *et al.* [105]	46 PAF + 53 healthy	2h	• Gaussian fit parameters (A,C, W) and error	Linear regression slope of the variability	Stepwise discriminant analysis (86.7 %)
Sovilj *et al.* [24]	14 POAF + 36 healthy	48h after CABG	• P-wave intervals • Heart rate • Wavelet energies and entropy	Cumulative rank with statistically significant features	Decision tree [applied over the time-course] (85.3 %)
Ovreiu *et al.* [106]	37 POAF + 53 healthy	30m	• Number of premature atrial complexes • HRV: mean, SDRR, rMSSD, total power, LF/HF, entropy • P-wave duration, amplitude, shape, inflection point, energy ratio	Mean value	Neuro-fuzzy (70.0 %)
Martinez *et al.* [107]	46 PAF	2h	• P-wave intervals • P-wave area, arc-length	Central tendency measurement	Decision tree (90.0 %)
Alcaraz *et al.* [108]	46 PAF + 53 healthy	2h	• P-wave frequency energies	Linear regression slope of the variability	Stepwise discriminant analysis (80.0 %)
Censi *et al.* [27]	73 pers. AF + 20 healthy	n/a	• P-wave duration and dispersion • P-wave morphology: polarity changes, fragmented conduction index • P-wave variability: amplitude dispersion index, correlation coefficient, DTW	Mean value	Receiver operating characteristic curves (without cross-validation) (100.0 %)

AF, atrial fibrillation; CABG, coronary artery bypass grafting; PAF, paroxysmal atrial fibrillation; POAF, postoperative atrial fibrillation; pers., persistent; n/a, not available; DTW, dynamic time warping.

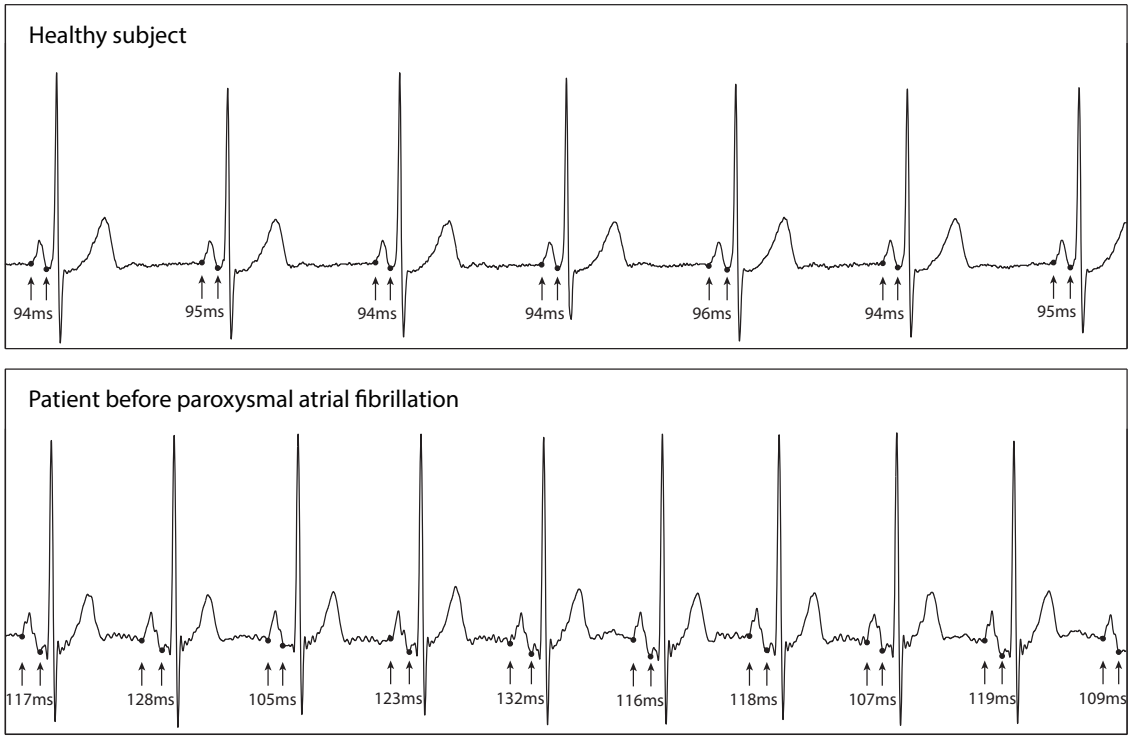

Figure 3.3: Example of representative ECG signals corresponding to healthy sujects (top) and patients susceptible to paroxysmal atrial fibrillation (PAF; bottom). Patients who will develop PAF are characterised with higher variability in P-wave duration (and other time and morphological features) than healthy subjects. Adapted from Alcaraz *et al.* [109].

unsupervised monitoring in the clinical setting. Specifically, at least three aspects need to be considered:

1. Systematic comparison of predictive methods. Although the majority of the proposed predictive techniques have achieved high accuracy, it is not clear which method, if any, is most suitable for clinical use. Moreover, it would be useful to investigate whether combining multiple indices provides improved prediction ability.

2. Evaluation of predictive methods in clinically useful scenarios. The majority of studies have studied the two hours before the onset of AF, which might be too late for prophylaxis. Furthermore, these studies have only used ECG signals during sinus rhythm, not having considered other rhythms that are common in the postoperative period, such as tachyarrhythmias or bradyarrhythmias [110]. Hence, it is necessary to test the prediction of POAF under clinically relevant and useful scenarios that mimic the real world clinical setting, and at a time before POAF where it is still possible to administer efficient prophylactic treatments.

3. Adaptation for unsupervised use during continuous monitoring. ECG data, especially when acquired during continuous monitoring, is susceptible to several artefacts that can lead to inaccurate measurements and consequently to a high number

of false alarms. This is particularly important in P-waves, which are highly susceptible to noise. Most studies have excluded noise-corrupted P-waves manually, which is not feasible in the clinical setting. Hence, the development of an automated P-wave quality assessment tool to discriminate between high quality and unreliable P-waves is needed to ensure the accuracy of techniques that predict AF.

Hence, several important aspects need to be considered so that the prediction of AF might be performed in the unsupervised clinical setting.

Chapter 4

P-wave Quality Index

This chapter presents and assesses the utility of the P-wave Quality Index (PQI), a tool which automatically identifies artefact-corrupted P-waves. Firstly, the PQI tool is described, and three different versions are proposed. The performance and utility of each of the three was assessed, and the version showing the highest performance and utility was selected to constitute the final PQI tool. Performance was assessed using its ability to discriminate between high and low quality data. Utility was assessed using: 1) mean absolute error comparisons of several well-known P-wave features with and without the PQI; and 2) evaluation of performance for predicting atrial fibrillation (AF) by using metrics extracted from the P-waves with and without the PQI.

4.1 Introduction

Long-term continuous monitoring using wearable devices has enabled early identification of several types of arrhythmia. A challenge to the use of long-term monitoring, where data is normally collected without clinical supervision, is ensuring that only high quality signals are used to derive clinical measurements. Physiological parameters extracted from artefact-corrupted signals may be inaccurate, which can lead to a high frequency of false alarms [111]. Therefore, assessment of signal quality is a crucial step before accurate and precise data analysis, such as extracting features, identifying deteriorations, and generating alerts.

The study of P-waves from surface electrocardiogram (ECG) is especially important in atrial arrhythmias, such as AF. For instance, subtle alterations in P-wave morphology have been found to be predictive of AF [23, 25–27]. The relatively low amplitude of the P-wave makes it highly susceptible to noise, severely affecting the extraction and quantification of its features, and hindering its clinical applications, especially in long-term recordings.

As artefact-corrupted P-waves influence derived measurements, it is essential to exclude them. This has been performed either manually, where expert cardiologists excluded unreliable P-waves by hand [25, 26], or automatically, with a conventional template-comparison method [112]. In the latter, individual P-waves were discarded if they had a cross-correlation coefficient lower than 0.7 with a template P-wave obtained with an averaging procedure [27]. Despite being promising, the methodology was not described in depth, and the threshold used to identify low quality P-waves was not optimized, limiting its reproducibility and utility.

Currently, even though much focus has been placed in the creation of methods that use P-waves to predict atrial arrhythmias, there is no reliable tool that assesses P-wave quality automatically. Thus, the clinical utility of such methods could be highly limited by parameters being estimated erroneously from low quality P-waves, causing false alerts. We propose a novel P-wave quality assessment tool to discriminate between high quality and unreliable P-waves, and assess its performance and utility on a database of wearable ECG recordings.

4.2 Methods

4.2.1 Data description

The Computers in Cardiology 2001 challenge database (AFPDB) [90, 113] was used to develop and assess the performance of the PQI [90]. This database was originally assembled with the aim of developing techniques for predicting paroxysmal AF (PAF). It consists of a learning and testing set, both of 100 records, with each record lasting 30 minutes. All signals have a sampling frequency of 128Hz and a 12-bit resolution. The database contains three different subject types: healthy controls, PAF patients far away from the arrhythmia onset (no episode 45 minutes before or after the record), and PAF patients just before the arrhythmia onset.

Forty four records (23 from controls and 21 from PAF patients) were randomly selected from the training select, after exclusion of those records which did not have discernible P-waves. After pre-processing with a bandpass filter with cut-off frequencies of 0.5 and 40Hz, the ECG lead in which P-waves were most visible for each record was chosen, and P-wave quality was manually annotated and double-checked. Following the tool's methodology (described bellow), P-waves were classified into three distinct classes, as illustrated in Figure 4.1.

Class A: High quality clean P-waves;

Class B: Complete noise and absent P-waves. This included P-waves with *no resemblance* to normal P-wave morphology, either due to motion artefacts, severe baseline wander, muscular activation interference, or simply the absence of a P-wave;

Class C: Unreliable, noise-distorted P-waves. This included P-waves that had *some resemblance* to normal P-wave morphology, but were still unreliable (i.e., their morphology was still excessively affected by noise). The degree of distortion varied from mildly to heavily distorted (see Figure 4.1).

This resulted in 22 hours of recording, corresponding to 97,989 P-waves: 88,606 class A (90.4%), 5,102 class B (5.2%), and 4,281 class C (4.4%). PAF patients had a higher proportion ($p<0.05$) of class B P-waves than controls, as compared using a two-sample *t*-test.

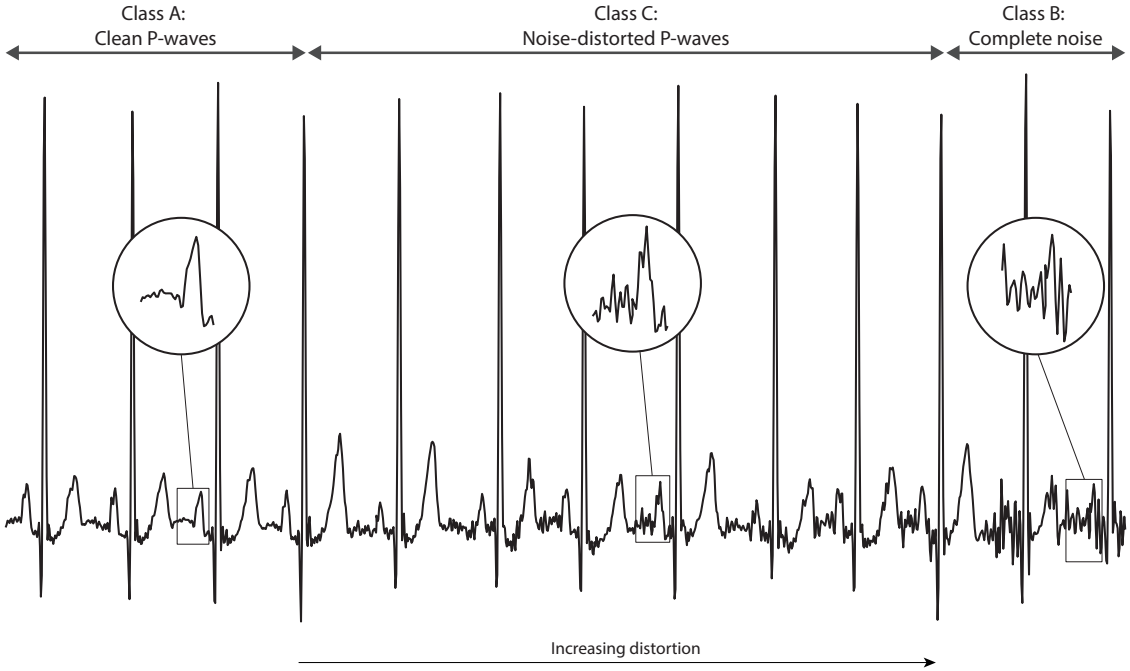

Figure 4.1: P-waves were classified into three different classes: high quality clean P-waves (class A), complete noise or absent P-waves (class B), and unreliable, noise-distorted P-waves (class C). Even though class C P-waves had some resemblance to normal P-wave morphology, they were still considered unreliable. For that class, the degree of distortion varied from mildly to heavily distorted.

4.2.2 P-wave Quality Index algorithm

The PQI tool algorithm is depicted in Figure 4.2. Briefly, the process started with P-wave detection and signal extraction, and was followed by two different decision-making stages based on template-comparisons. The first decision stage aimed to remove P-waves with no resemblance to the normal P-wave morphology (class B), while the second was more refined, removing P-waves still excessively distorted by noise, and hence unreliable (class C).

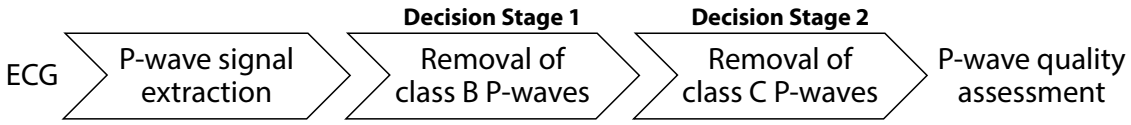

Figure 4.2: The P-wave quality index (PQI) tool consisted of three steps. In the first, the P-waves were identified and extracted. These signals were then used to create P-wave templates and to extract features, which were used in two different decision-making stages. At the end of the PQI, P-waves distorted by noise, and hence unreliable, were expected to be excluded from the analysis.

4.2.2.1 P-wave signal extraction

The first step towards assessing the P-waves' quality was to detect them in the ECG and extract them. This signal extraction step was of high importance, as the following PQI algorithm stages relied on the accurate segmentation of the P-waves. Hence, a proper P-wave quality assessment was only possible with a proper P-wave signal extraction. In order to extract the P-wave, its peak was first identified. To do so, first, each R-peak was identified using the widely used Pan, Hamilton and Tompkins algorithm [100, 101], followed by Q-wave and P-wave peak delineation using the phasor transform delineation algorithm [104] (see Section 5.2.2 for detailed explanation).

Finally, P-wave signals were then extracted by taking a window of width 300ms, centred on the correspondent P-wave peak. This window length ensured that the whole P-wave plus its surrounding baseline signal were captured. Any windows containing the correspondent Q-wave were shortened to just before that point.

4.2.2.2 Decision stage 1: Removal of completely noisy P-waves

The aim of the first decision-making stage was to exclude P-waves which were heavily distorted, either due to motion artefacts, severe baseline wander, muscular activation interference, or simply by the absence of a P-wave (class B). This first stage had the additional purpose of removing P-waves that could influence the more refined templates created during the second decision-making stage.

P-waves within a 30-minute period were aligned, and a P-wave template was created in an iterative process. Before the alignment process, differences in P-wave signal durations due to different PQ intervals and/or imperfect Q-wave delineations were accounted for by further shortening the windows' end so that it was not exceeded by more than 30% of the individual windows' end.

The alignment process allowed each individual P-wave signal to be aligned with a template. This template started as the second P-wave, and was updated for every subsequent P-wave. Hence, the template used to align a given P-wave corresponded to the average of all the previous aligned P-waves. The optimum alignment was that which maximized the Pearson correlation coefficient between the P-wave and template, being allowed a maximum shift from the original position of 50ms. Given that signals were always pre-aligned

by their peak, this maximum shift aimed to ensure that only physiological variations were present, averting unnecessary overfitting with noise. As illustrated in Figure 4.3, this alignment process was of utmost importance, especially for fragmented P-waves.

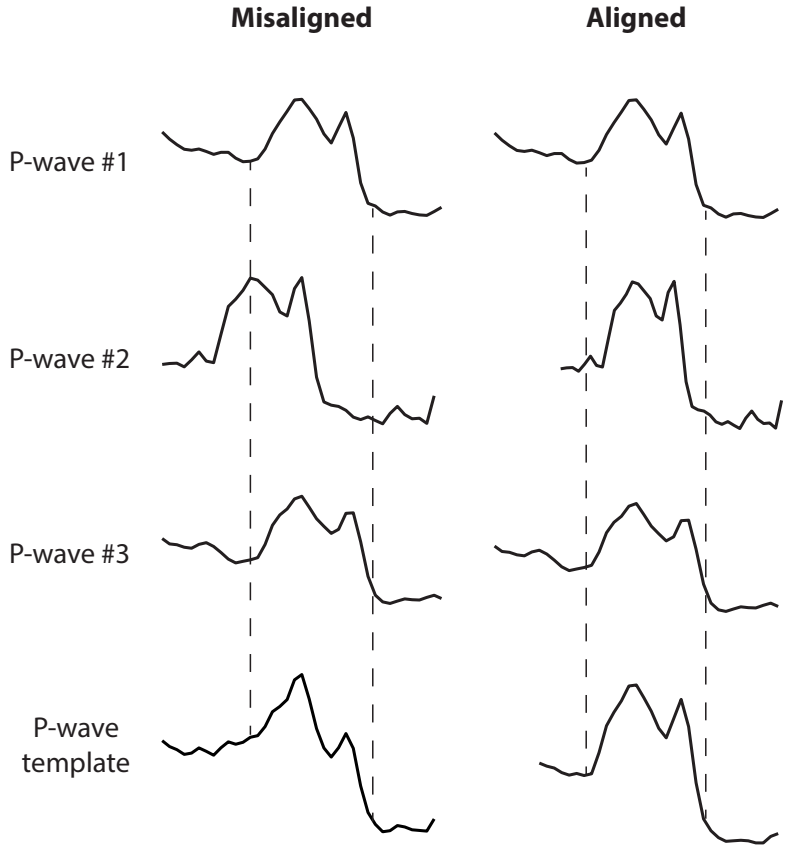

Figure 4.3: Proper P-wave alignment was essential for an optimal P-wave template creation. When P-waves are aligned by their peak (left), the correspondent created template can fail to represent the associated P-wave morphology. This is due to variations in P-wave peak identification, which can cause P-waves to be misaligned. As depicted, this effect is especially dangerous in fragmented P-waves where P-wave peak delineation varies greatly. However, with proper P-wave alignment (right), the template creation is optimized.

After all the P-waves were aligned with the corresponding template, several features were extracted in order to retrieve information about them. These features were then used as candidate features for the PQI to assess P-wave quality. The following features were tested:

1. P-wave template comparison features:

    a) Pearson's correlation coefficient between P-wave and template [114];

    b) Root mean square (RMS) difference between P-wave and template;

    c) Pearson's correlation coefficient between P-wave and template derivatives;

    d) Root mean square difference between P-wave and template derivatives.

2. P-wave signal features:

    a) Standard deviation of P-wave [115];

    b) Number of zero-crossings of P-wave [115];

    c) Area of P-wave;

    d) Kurtosis of P-wave [111];

    e) Skewness of P-wave [111];

    f) Standard deviation of P-wave derivative.

The ability of the aforementioned features to discriminate between class A and class B P-waves was assessed through comparisons with the manual annotations. Performance was quantified using the receiver operating characteristic (ROC) curve, and features with an area under the curve (AUC) greater than 0.75 were selected be included in the final decision stage 1 model. This model consisted of a decision tree and its performance was assessed using 10-fold cross validation. During the analysis, class A clean P-waves were considered as the positive class.

4.2.2.3 Decision stage 2: Removal of distorted P-waves

Having removed the completely distorted P-waves, more refined P-wave templates could be created. This was then used in a final decision stage with the aim of removing distorted and unreliable P-waves (class C), while being able to accommodate possible P-wave morphological variations.

This second decision stage was similar to the first, but with the single difference that P-waves were aligned and a template created every few (n) P-waves (instead of every 30 minutes), aiming to accommodate both ECG physiological variations along time (Figure 4.4) and the greater P-wave variability that precedes AF [25–27]. The value of n was varied from 10 to 100 and the number which gave the best overall feature results was selected and used to build the final model. Similarly to the first decision-making stage, the same features were extracted and their selection process used the same criteria. Finally, two decision trees were assembled - one promoting balanced (i.e., assigning equal importance) sensitivity and specificity, and the other favouring high sensitivity.

4.2.3 P-wave Quality Index performance and utility assessment

Once the PQI decision-making models were built, three different versions of the PQI tool were assembled:

PQI Stage 1: Using *only* Decision Stage 1;

Balanced PQI: Using Decision stage 1, plus Decision Stage 2 with balanced sensitivity and specificity;

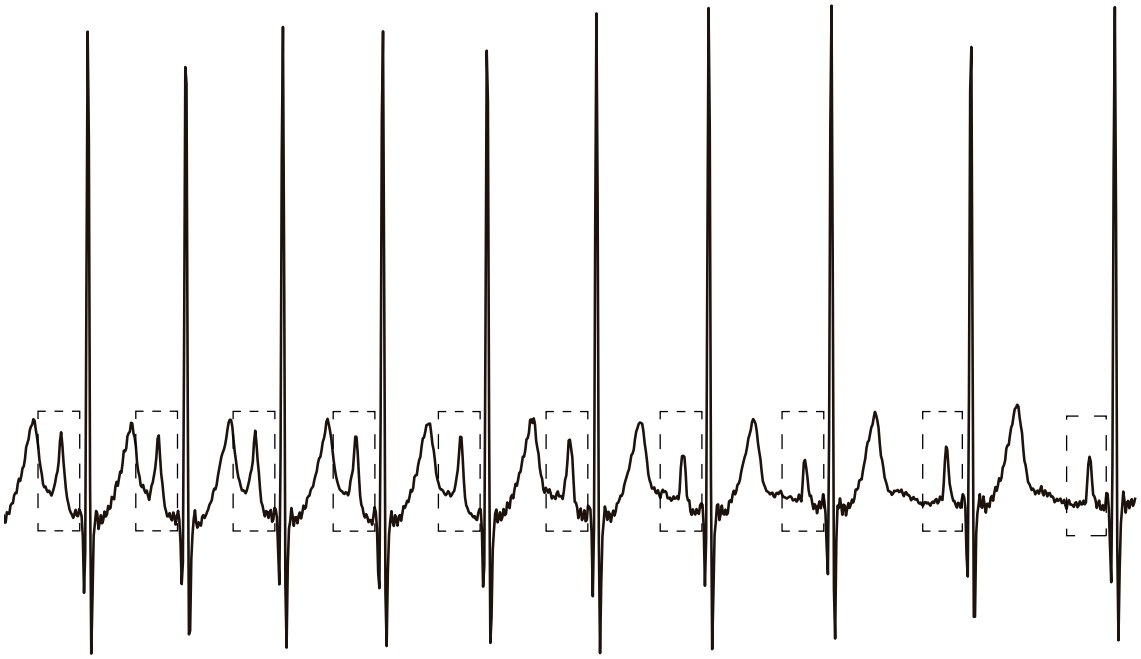

Figure 4.4: P-wave morphology variation over time due to physiological variations. The extracted P-wave signal (dotted box) starts by being asymmetrical and containing the T-wave from the previous beat, and ends up symmetrical and containing the surrounding baseline.

Sensitive PQI: Using Decision stage 1, plus Decision stage 2 with high sensitivity.

In order to select the best suited version of the PQI tool, the classification performance and two aspects of their utility were compared:

1. Classification performance (i.e., the ability to distinguish between high and low quality P-waves) assessment of each PQI version on the entire 44 records. Class A clean P-waves were again considered as the positive class, while classes B and C were merged into one unique negative class.

2. Mean absolute error (MAE) comparisons between P-wave features obtained from all the recorded P-waves, and only from those classified as high quality by each PQI tool version. Clinical features included the mean and standard deviation of P-wave duration, amplitude, area and arc-length, as they have shown to give important insights on atrial conduction [23, 26]. Reference values of P-wave features were obtained from P-waves which were manually annotated as high quality.

3. Evaluation of performance for predicting PAF. Several features were extracted from the P-waves (see Table 5.1 from Chapter 5), resulting in feature time-series, from which a selection of metrics were calculated (see Table 5.2). Next, statistical significance (Mann-Whitney) tests were used to evaluate the ability of the extracted metrics to distinguish between the three different subject groups (Healthy, PAF far,

and PAF close), resulting in each metric being identified as either significantly different between groups, or not significantly different. This process was performed four times: using the manually annotated high quality P-waves, and using the P-waves labelled as high quality by each of the three versions of the PQI tool. An evaluation of performance for predicting PAF was performed for each version of the PQI by testing if the set of significant metrics obtained was similar to the set obtained when using the manual annotations. This was performed using the sensitivity and specificity statistics, which were calculated by comparing the significance of the metrics obtained from each version of the PQI tool with the significance of the metrics obtained from the manually annotated P-waves. Significant metrics ($p<0.05$) were considered the positive class, while non-significant metrics ($p\geq0.05$) constituted the negative class. Tested metrics included the mean, standard deviation (both on entire records and on a 5-minute basis), and variability (linear and non-linear) measurements of several P-wave features (refer to Section 5.2.4 for more details), which have been shown to be predictive of AF.

Once the best PQI version was selected, two additional quality assessments were performed. Firstly, it was tested whether the P-wave quality classification performance of the final PQI tool differed between controls and PAF patients. This was performed by comparing the individual classification performance of the two groups using a two-sample *t*-test. Secondly, the ability of the final PQI tool to separately distinguish class A from class B (class A *vs.* class B) and class A from class C (class A *vs.* class C) P-waves was assessed on the entire 44 records.

Throughout all the mentioned P-wave quality classification performance assessments, only the sensitivity and specificity metrics were used, as they have the advantage of being independent to class distributions [116, 117], and therefore are able to provide comprehensive assessment of imbalanced learning problems [118], such as the present one.

4.3 Results

4.3.1 Decision stage 1: Removal of completely noisy P-waves

Table 4.1 shows the performance of each separate feature, measured using the AUC, on discriminating class A from class B P-waves. The following features had AUCs greater than 0.75, and were therefore used to estimate the decision-making stage 1 model: template correlation, template RMS error, template derivative correlation, and P-wave skewness.

A decision tree was built using the aforementioned features. Given the high imbalance between classes, this model was built assuming both classes' probability as equal, allowing one to balance sensitivity and specificity. Furthermore, the decision tree had a maximum of 5 splits (i.e. decisions), in order to avoid overfitting. Figure 4.5 shows the obtained decision tree, and Table 4.2 shows its performance measured using 10-fold cross

Table 4.1: P-wave quality assessment features, and corresponding area under the curve (AUC) for both decision-making stages. Decision stages 1 and 2 aimed to distinguish class A from class B P-waves, and class A from class C P-waves, respectively.

	Area under the curve	
	Decision stage 1	**Decision stage 2**
P-wave template comparison features		
Template correlation	**0.96**	**0.89**
Template RMS difference	**0.96**	**0.85**
Template derivative correlation	**0.92**	**0.83**
Template derivative RMS difference	0.73	0.71
P-wave signal features		
Standard deviation	0.67	0.69
Zero crossing number	0.56	0.62
Area	0.60	0.54
Kurtosis	0.61	0.57
Skewness	**0.76**	0.66
Derivative standard deviation	0.64	**0.78**

RMS, root mean square.

Table 4.2: Performance of decision-making stage 1 on distinguishing class A from class B P-waves, and performance of each version of decision stage 2 to distinguish class A from class C P-waves.

	Sensitivity	**Specificity**
Decision Stage 1	0.94	0.96
Decision Stage 2		
Balanced version	0.83	0.89
Sensitive version	0.99	0.22

validation. Note that, even though the skewness feature had shown good discrimination ability, it was not included in the decision tree.

4.3.2 Decision stage 2: Removal of distorted P-waves

Several P-wave groupings were tested in order to maximise discernment between class A and class C P-waves. Keeping in mind that increasing window duration (n) has the disadvantage of increasing time between consecutive windows, due to gaps in the data [119], and balancing with the obtained AUCs (Figure 4.6), n was set to be 20 P-waves.

When distinguishing between class A and class C, the features template correlation, template RMS error, derivative template correlation and P-wave derivative standard deviation had AUCs greater than 0.75 (Table 4.1). Two different decision trees were built using those features, both with a maximum of 5 splits (Table 4.2): one favoured balanced

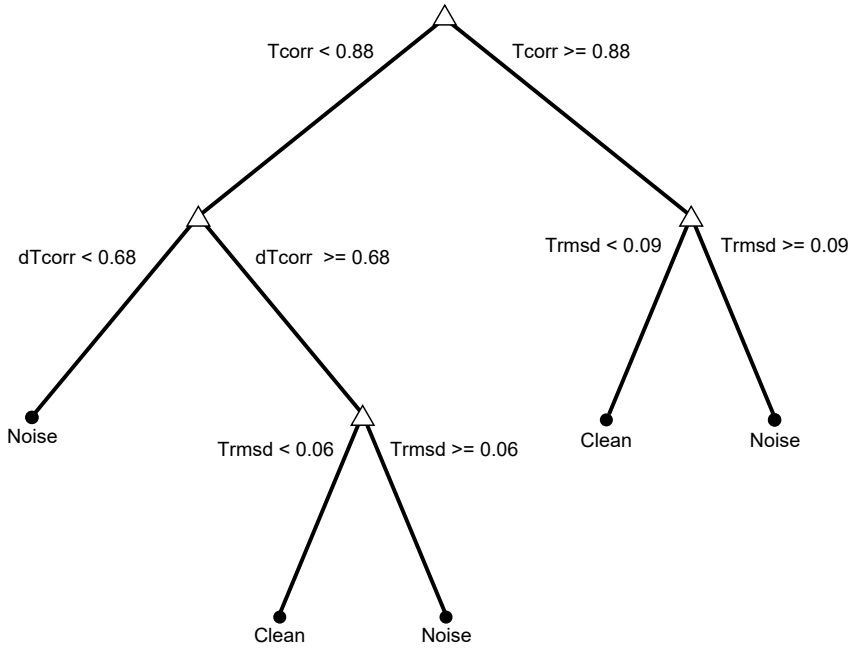

Figure 4.5: Final decision-making stage 1 decision tree. This model made use of the following features to discern clean from complete noise/absent P-waves: template correlation (Tcorr), template root mean square difference (Trmsd), and template derivative correlation (dTcorr).

sensitivity and specificity and was used in the Balanced PQI version, while the other promoted high sensitivity (Figure 4.7) and was used in the Sensitive PQI version.

4.3.3 P-wave Quality Index performance and utility assessment

After the three different versions of the PQI tool (PQI Stage 1, Balanced PQI, and Sensitive PQI) were assembled, their classification performance and utility were assessed in order to select the best suited version:

1. Classification performance assessment (Table 4.3). Overall, the PQI Stage 1 and Sensitive PQI versions had high sensitivity and good specificity, while the Balanced PQI had high specificity and good sensitivity. The Sensitive PQI version had the best sensitivity/ specificity relationship.

Table 4.3: Classification performance (i.e., ability to distinguish between high quality (class A) and low quality (classes B and C) P-waves) of each version of the PQI tool.

	Sensitivity	**Specificity**
PQI Stage 1	0.93	0.77
Balanced PQI	0.77	0.98
Sensitive PQI	0.93	0.82

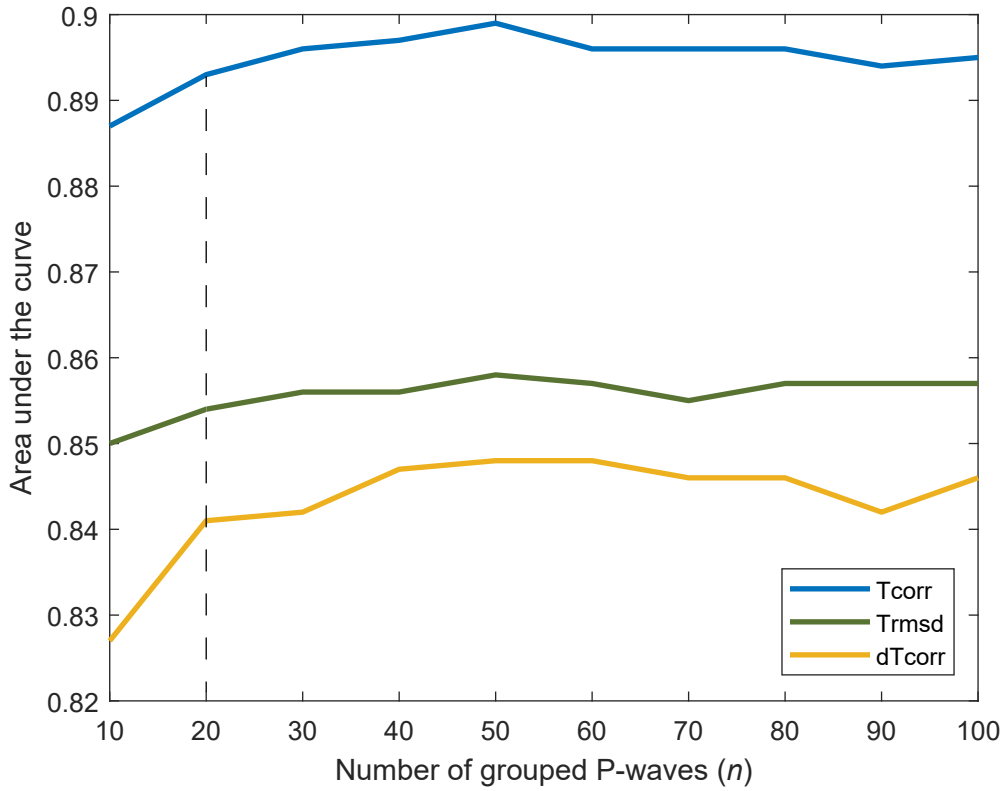

Figure 4.6: Area under the curve (AUC) in function of number of P-waves grouped in order to create a P-wave template, n. The value of n was set to 20 P-waves by balancing between high AUC and window length. The features template correlation (Tcorr), template root mean square difference (Trmsd), and template derivative correlation (dTcorr) are here represented given their AUC greater than 0.75.

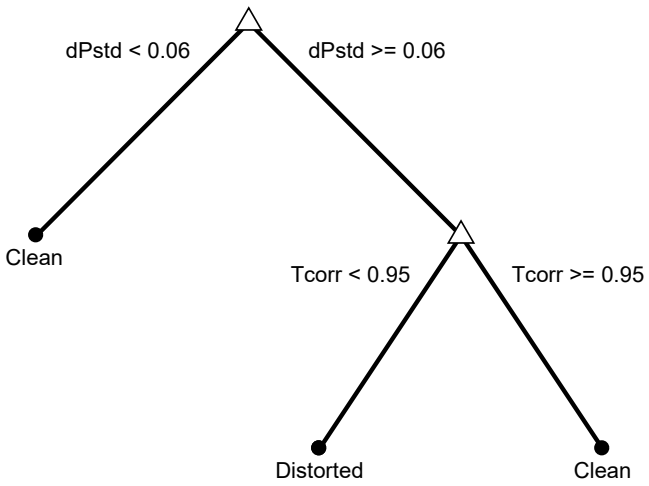

Figure 4.7: Final decision-making stage 2 decision tree, correspondent to the Sensitive PQI tool version. This model made use of the following features to distinguish clean and noise-distorted P-waves: standard deviation of P-wave derivative (dPstd) and P-wave template correlation (Tcorr).

2. MAE comparisons (Table 4.4). Overall, the Sensitive PQI version outperformed the remaining by having the smallest MAE, while the Balanced PQI had the highest MAE across the majority of the tested metrics. Figure 4.8 shows, for each record, the difference in MAE for the mean P-wave duration using the Sensitive PQI version (i.e. the MAE obtained using the P-waves obtained with Sensitive PQI subtracted from the MAE using all P-waves), representing a case where the Sensitive PQI brought an intermediate improvement on feature assessment.

3. PAF prediction analysis comparisons (Table 4.5). The Sensitive PQI version outperformed the remaining in terms of sensitivity/ specificity relationship throughout most of the tested methods. Furthermore, the analysis using all the P-waves contained in the recordings was found to have the greatest total sensitivity, driven by the non-linear variability method.

Table 4.4: Utility of each version of the P-wave quality index (PQI) tool assessed with mean absolute error (MAE) comparisons between P-wave features obtained using all the available P-waves and those labelled as high quality by each version of the PQI tool. The presented factor represents how many times the given method reduced the MAE in comparison to when using all the P-waves.

		Mean Absolute Error						
		All	**PQI Stage 1**		**Balanced PQI**		**Sensitive PQI**	
P-wave features	**Metrics**	**Error**	**Error**	**Factor**	**Error**	**Factor**	**Error**	**Factor**
Duration [ms]	**Mean**	2.63	0.55	4.8	1.04	2.5	0.50	**5.3**
	SD	3.26	0.39	8.4	0.68	4.8	0.34	**9.4**
Amplitude [mV]	**Mean**	5.73	1.85	3.1	0.28	2.0	0.17	**3.4**
	SD	0.24	0.03	**7.2**	0.05	5.3	0.04	7.0
Area [μVs]	**Mean**	0.04	0.01	3.2	0.02	2.5	0.01	**3.3**
	SD	0.19	0.02	**8.1**	0.03	6.0	0.02	7.9
Arc-length [μVs]	**Mean**	0.34	0.07	4.8	0.13	2.5	0.06	**5.3**
	SD	0.42	0.05	8.4	0.09	4.8	0.04	**9.3**

SD, standard deviation.

Given that the Sensitive PQI version outperformed the others in classification performance and in both utility assessments, it was chosen to constitute the final PQI tool and will be, from now on, referred only as "PQI".

Finally, the ability of the final PQI tool to separately distinguish class A from class B (class A *vs.* class B) and class A from class C (class A *vs.* class C) P-waves was assessed on the entire 44 records, and overall performance between controls and PAF patients was compared (Table 4.6). Class B P-waves were removed with extremely high specificity (0.97), while class C P-waves were removed with low specificity (0.64). Furthermore, no statistically significant difference ($p<0.05$) was found in sensitivity nor specificity between the PAF and control groups, even though the controls had a lower mean specificity value. When inspecting performance in an inter-individual fashion, 4 records (3 from

Table 4.5: Utility of each version of the P-wave quality index (PQI) tool assessed by evaluating their performance for predicting paroxysmal atrial fibrillation (PAF). This was performed by testing if the significance of the computed predictive metrics was correct, using the sensitivity and specificity statistics. The metrics used to predict PAF are explained in detail in Section 5.2.4.

	Sensitivity/ Specificity			
Method	**All**	**PQI Stage 1**	**Balanced PQI**	**Sensitive PQI**
Feature's mean, SD, and time course	0.64/ 0.96	0.82/ 0.99	0.80/ 0.98	0.86/ 0.99
Linear variability and time course	0.31/ 0.78	0.70/ 0.92	0.52/ 0.94	0.68/ 0.92
Non-linear variability (CTM)	0.31/ 0.74	0.22/ 0.92	0.11/ 0.68	0.24/ 0.92
Total	0.22/ 0.63	0.16 /0.88	0.08/ 0.59	0.18/ 0.88

CTM, central tendency measurement; SD, standard deviation.

controls and 1 from a PAF patient) had a specificity lower than 0.5 (Figure 4.9). From those, one had a null specificity due to the existence of only two noisy P-waves on the entire recording, another also had a very low percentage of low-quality P-waves, and the remaining two were constituted mostly by class C P-waves with a low level of distortion.

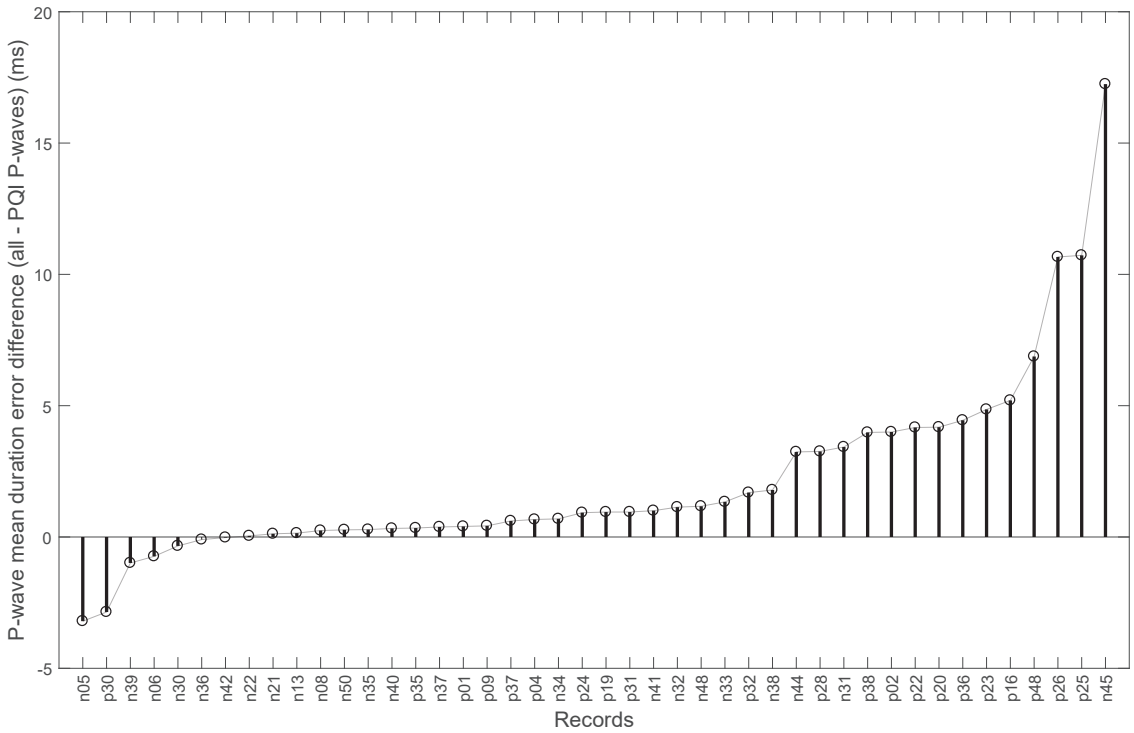

Figure 4.8: Differences in mean absolute error (MAE) between P-wave mean duration calculated when using all the P-waves and when using those obtained with the final P-wave quality index (PQI) tool (correspondent to the Sensitive PQI version). A positive difference indicates an improvement of MAE when using the proposed PQI tool.

4.4 Discussion

The future implementation of techniques that make use of the P-wave to predict atrial arrhythmias, such as AF, will highly rely on proper P-wave quality assessment. Currently there is no reliable tool that allows to discard noise-corrupted P-waves, despite its importance. In the present chapter we proposed the PQI, a novel tool capable of automatically identifying noise-corrupted, and hence unreliable, P-waves.

Briefly, the algorithm starts by detecting and extracting the P-waves' signal, which is then used in two decision-making stages: the first with the aim of removing highly noisy or absent P-waves, and the second with the aim of removing less distorted, but still unreliable, P-waves. We have shown the PQI tool to have a positive clinical impact by reducing the error of commonly used P-wave features and by improving the ability to predict AF. Finally, we have shown the PQI to perform equally well on healthy subjects and patients who will develop PAF, further indicating that the proposed tool was able to accommodate the P-wave variability that precedes AF [25–27].

Table 4.6: Comparison between individual P-wave quality classes (class A *vs.* Class B and Class A *vs.* Class C P-waves), and comparison of overall classification performance (class A *vs.* classes B and C) between controls and patients who developed atrial fibrillation.

	Comparison between P-wave quality classes		**Comparison of overall performance between subject groups**			
Statistic	**Class A *vs.* Class B**	**Class A *vs.* Class C**	**Total**	**Controls**	**AF patients**	***P*-value**
Sensitivity	0.93	0.93	0.93	0.92	0.95	0.33
Specificity	0.97	0.64	0.82	0.76	0.86	0.08

AF, atrial fibrillation; class A, clean P-waves; class B, complete noise/ absent P-waves; class C, unreliable, noise-distorted P-waves.

The proposed tool was built using a large dataset containing almost 100,000 manually annotated P-waves (corresponding to 22 hours of recording) across several different morphologies, and trialled 10 different P-wave quality assessment features, some of them novel. In addition, the proposed tool was built using simple decision tree models, making it more likely to work in other environments than the current one.

In this study, three different versions of the PQI tool were proposed: one using only the first decision stage (PQI Stage 1), and two other using the complete PQI tool, but using the second decision stage with balanced sensitivity and specificity (Balanced PQI) and high sensitivity (Sensitive PQI). The Sensitive PQI version was found to be the best performing (in terms of utility), followed by the PQI Decision Stage 1 and the Balanced PQI. Even though the Balanced PQI was the version with the highest P-wave quality classification specificity (0.98), it ended up being the least useful one. This can be explained by the not-so-high sensitivity (0.77): clean P-waves with different characteristics

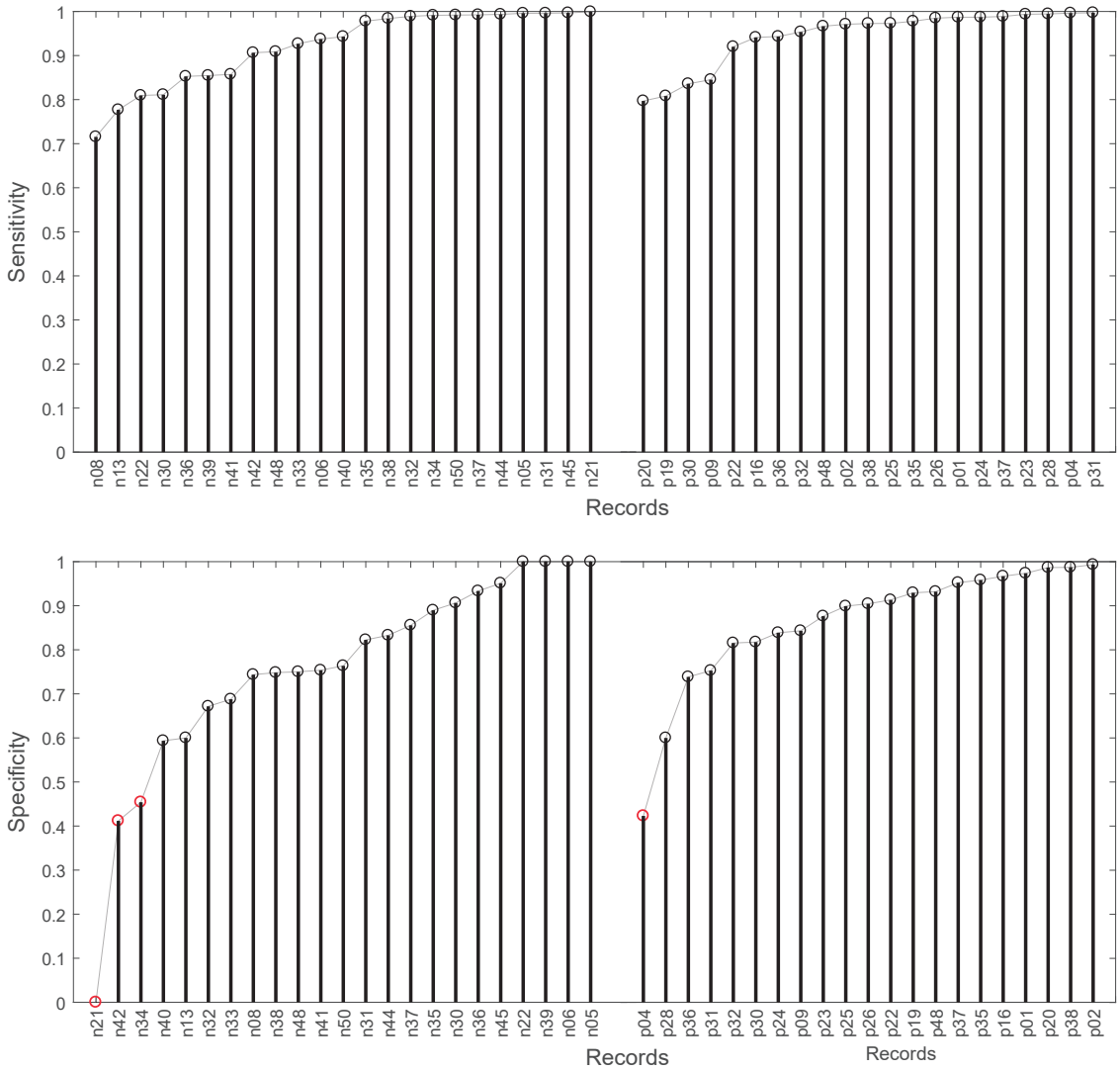

Figure 4.9: Individual performance to classify overall P-wave quality (class A *vs.* classes B and C), using sensitivity (top) and specificity (bottom), of the P-wave quality index (PQI) tool on controls (left) and patients (right). Records with performance lower than 0.5 are highlighted in red.

(e.g. duration) ended up being removed (especially in the more rigorous second decision stage), prejudicing both the MAE and PAF prediction analyses, as those P-waves contained important and useful information.

During the PAF prediction analyses, the analysis conducted using all the P-waves from the recordings was found, surprisingly, to have the best PAF-prediction sensitivity (Table 4.5). Since this result was driven by the non-linear variability metrics, this is likely explained by the higher proportion of complete noise/absent P-waves present in PAF patients, given that features extracted from noisy P-waves are associated to higher variability. Nonetheless, the specificity value obtained for this particular analysis was low, reflecting a clear example where noisy P-waves affect data analysis by leading to a high number of false alarms.

It is important to note that, even though the Sensitive PQI version (i.e. the final PQI tool) had lower specificity in the second decision stage, the majority of the more-distorted class C P-waves were removed during both decision stages, while mildly-distorted (but still labelled as unreliable) class C P-waves were retained (hence the overall 0.64 specificity; Table 4.6). Given that the more-distorted class C P-waves were more different from the clean P-waves than the mildly-distorted class C P-waves, they were more easily identified. Hence, the remaining class C P-waves that were labelled as high quality (false positives) probably did not affect the analysis excessively.

A way in which specificity could be increased in the future is with the previous characterization of the signal being analysed in terms of either signal-to-noise ratio (SNR) or lead configuration, hence allowing to adapt the decision rules to specific signals. In addition, it is also important to note that the P-wave's quality annotation procedure was associated to some subjectivity while separating clean from mildly, yet considered unreliable, P-waves. Nonetheless, as noted, in the proposed tool, those P-waves were likely labelled as of high quality.

Finally, the proposed tool can work on a near real-time basis, an important feature of quality assessment tools aimed to be used during continuous monitoring. Even though the PQI requires 30-minute signal portions, such time precision is enough to predict and act upon atrial arrhythmias such as AF.

4.4.1 Limitations and Future Work

4.4.1.1 P-wave signal extraction and template creation

The proposed tool assumes that, during the creation of a P-wave template each 30 minutes, noise is cancelled out and, therefore, the obtained signal reflects a clean P-wave. However, during long term continuous monitoring this may not be the case, as there is the possibility that the recording may be highly corrupted by noise during a good portion of time (e.g. during poor electrode contact with the skin), thus causing the obtained P-wave template to be distorted, and hence affecting severely the tool's performance. This can be safeguarded in future works with the addition of a template-verification stage where, for example, the obtained template signal could be compared with a Gaussian function. Another important case where the P-wave template might be distorted is during AF, as there is an absence of P-waves. This particular issue can be addressed in the future with the use of AF arrhythmia detection algorithms, which would then trigger a modification in the template creation procedure.

Furthermore, during the creation of template P-waves, it is assumed that the signal has a single P-wave morphology, thus ignoring the possibility of a secondary P-wave morphology. Recently, a study has found that an increased number of P-waves matching a distinct secondary morphology was significantly associated with PAF, possibly indicating different conduction routes on the atrial myocardium [120]. If those secondary morphologies significantly differ from the main morphology of the created template, they will

probably be classified as noise and hence discarded from the analysis. Thus, future improvements on P-wave quality assessment could include the possibility of an additional P-wave morphology with, for example, the creation of two different P-wave templates that could then be used to extract P-wave features and assess its quality.

4.5 Final Remarks

In this chapter we presented the P-wave Quality Index - a novel tool capable of automatically rejecting noise-corrupted P-waves - and showed its positive impact and utility on P-wave feature assessment and AF prediction. This tool will be particularly important with the future clinical implementation of AF prediction methods, since the reliability of P-wave measurements is of utmost importance for their accuracy.

The final proposed tool was able to exclude artefact-corrupted P-waves with a high sensitivity (0.93) and good specificity (0.82). Furthermore, the tool allowed for more precise measurements of several well-known P-wave features, and improved the prediction of AF, especially when using P-wave variability methods.

The P-wave Quality Index tool Matlab® code and P-wave's quality assessment dataset are being made publicly available at **http://github.com/diogotecelao/PwaveQualityIndex**. These resources allow future researchers to apply the tool in their studies, reproduce the analyses herein presented, and to use the dataset for additional studies.

Chapter 5

Prediction of Paroxysmal Atrial Fibrillation: Computers in Cardiology 2001 Challenge

This chapter presents a study on the prediction of paroxysmal atrial fibrillation (PAF). The aim of this study was to identify subjects at risk of developing PAF, and to predict the imminent PAF onset in patients known to be at risk. Several electrocardiogram (ECG) features and metrics, mostly derived from the P-wave, were extracted. The P-wave Quality Index tool (PQI; presented in Chapter 4) was used to automatically eliminate low-quality P-waves from the analysis. Finally, the ability of the extracted indices to predict PAF was studied. In this study, a selection of previously proposed methods were compared, and it was tested if their combination improved prediction of AF.

5.1 Introduction

Atrial fibrillation (AF) is the most common cardiac arrhythmia disorder [1], with its incidence expected to increase exponentially [11, 32, 33]. The most serious complications of AF include stroke, thromboembolic events, and heart failure. Furthermore, AF leads to electrophysiological modifications within the atria, such as electrical, contractile, and structural remodeling, which reduces the probability of AF termination [26], and favours the progression of PAF into persistent AF [121].

The development of accurate predictors of AF is clinically important because preventive as well as more effective treatments may by used to avert the loss of sinus rhythm. Consequently, the possibility of electrical stabilization and non-recurrence of AF increases greatly [25]. Furthermore, maintenance of sinus rhythm can reduce patients' risks and symptoms, improve hemodynamics, and decrease the possibility of recurrence

of AF [25].

The 2001 Computers in Cardiology (CinC) challenge had two major goals: to identify subjects at risk of developing PAF (event 1), and to predict the imminent PAF onset in patients known to be at risk (event 2). This challenge has stimulated research on the prediction of PAF by making a database with the pertinent electrocardiogram (ECG) recordings publicly available, and has provided a way in which researchers can benchmark their methods.

Intensive efforts have been made in the last decade to identify predictors of AF from the ECG. Several ECG features and metrics have been proposed. However, further investigation of these methodologies is required to determine their clinical utility. Furthermore, a systematic comparison of the methods has not been performed. Moreover, the combination of several indices may provide improved prediction performance. Finally, previously proposed methods were assessed using high quality ECG signals, and have not been assessed in an unsupervised fashion.

In this chapter we present a study on the prediction of PAF, in which the 2001 CinC challenge was followed. The aim of this study was to assess the performance of several methodologies to predict PAF when applied to an unsupervised clinical setting, along with the previously proposed PQI tool. This study is the first performing a systematic comparison of methods, including an investigation of which are more suitable for clinical use. This included the extraction of 43 features from the ECG, and the calculation of 21 different metrics. Finally, this study assessed the performance of those methods mostly in an unsupervised manner.

5.2 Methods and materials

5.2.1 Study population

The PAF prediction challenge database (AFPDB) [90, 113] consists a pair of 30-minute excerpts of two-channel 24h long-term ECG recordings from subjects falling into three different categories: healthy controls ("Healthy"), PAF patients far away from the arrhythmia onset (no episode 45 minutes before or after the record; "PAF far"), and PAF patients just before the arrhythmia onset ("PAF close"). For the healthy controls, the pair of ECG segments was chosen at random times in the 24 hour recording. The signals have a sampling frequency of 128Hz and a 12-bit resolution. The database is divided into a learning set and a testing set of equal size, and both contain approximately the same number of healthy subjects and patients who will develop PAF. The training set contains 50 records from healthy subjects, 25 from PAF far, and 25 from PAF close.

Given that the ECG recordings were not recorded with the special intention of studying the P-wave, 39 recordings from the training set were removed from the analysis due to the absence of discernible P-waves. This resulted in 61 records in the training set: 32 healthy, 14 PAF far, and 15 PAF close. Finally, for each subject, the channel in which

the P-waves were most pronounced was chosen and used throughout the analysis. The testing set was not used since no information is provided on which subjects belong into which category in this set.

5.2.2 Signal preprocessing and delineation

The original ECG signals were preprocessed by applying a forward/backward filtering strategy, an approach which has been demonstrated to preserve the P-wave fiducial points [104]. This filter consisted of a 2858^{th}-order bandpass filter with cut-off frequencies of 0.5 and 40Hz, and had the aim of removing baseline wander and high frequency noise.

5.2.2.1 R-peak detection

The delineation process started with R-peak detection using the widely used Pan, Hamilton and Tompkins algorithm [100, 101]. Briefly, the ECG signal was differentiated, squared, and a moving average filter was applied (Figure 5.1). Next, windows containing R-peaks (starting at QRS_{left} and ending at $\text{QRS}_{\text{right}}$) were defined using a signal-dependent threshold. Finally, the R-peak was identified as the maximum absolute value of the ECG signal within those windows.

5.2.2.2 Q-wave detection

Q-wave detection and P-wave delineation were performed based on the phasor transform (PT) [104]. The PT converted each instantaneous sample of the ECG signal (S[n]) into a phasor of amplitude M[n] and phase $\phi[n]$, such that:

$$PT\{S[n]\} = M[n]e^{j\varphi[n]} \tag{5.1}$$

where the magnitude M[n] and phase $\varphi[n]$ of this phasor were computed as:

$$M[n] = \sqrt{R_v^2 + S[n]^2} \tag{5.2}$$

and

$$\varphi[n] = \arctan\left(\frac{S[n]}{R_v}\right) \tag{5.3}$$

In this way, by considering the instantaneous phase variation $\varphi[n]$, the subtle alterations provoked by the ECG waves in the recording were maximised, regardless of eventual low amplitude, thus making their detection and delineation considerably easier (Figure 5.1).

The value of R_v determined the degree to which the ECG waves were enhanced by the PT: the lower the value of R_v, the higher the differences among phase variations in the complex plane [104]. Hence, R_v was changed according to the ECG wave aimed to be detected or delineated, as described further below.

The detected R-peaks served as a reference for the identification of the Q-waves. To do so, firstly, signal was extracted from an window beginning 50ms before QRS_{left} and

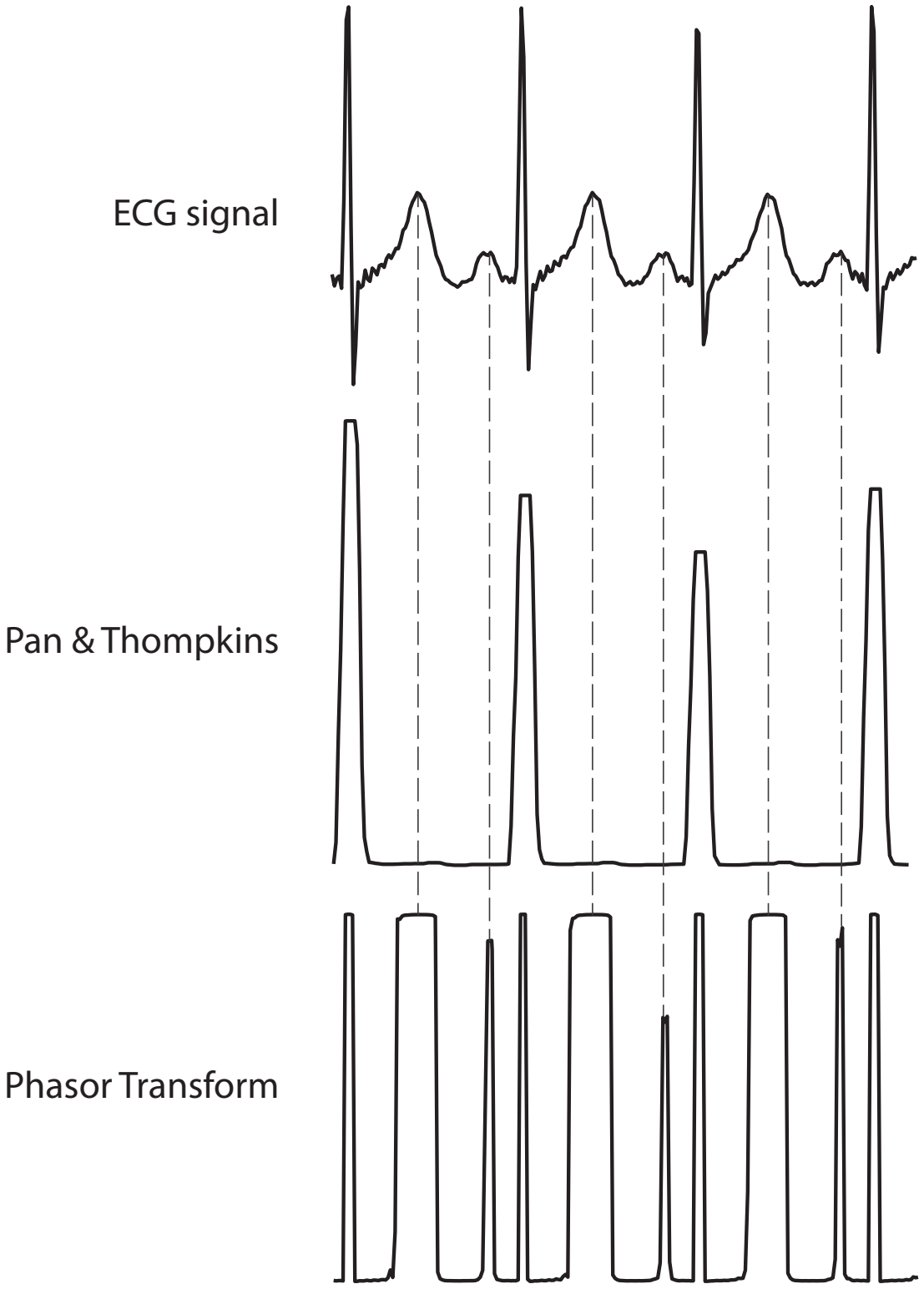

Figure 5.1: Transformations of the electrocardiogram (ECG) signal (top) obtained using the Pan & Thompkins algorithm [100, 101] (i.e. by differentiating, squaring, and applying a moving average filter; middle), and the phasor transform [104] (bottom). These transformations were used to delineate the ECG fiducial points.

ending at $\text{QRS}_{\text{right}}$ (of the corresponding R-peaks). Then, the absolute value of that portion of signal was calculated, its median was removed, and the PT was applied with $R_v = 0.001$. This approach was distinct from that originally described in the literature [104], where the PT is applied directly to the absolute value of the entire ECG signal, which causes $\varphi[n]$ to be strongly influenced by noise. Next, a boundary point γ_{QRS-} was defined as the closest point before the R-peak in which $\varphi[n]$ was lower than 25% of the maximum phase variation ($\pi/2$).

Afterwards, another signal portion was extracted with a window starting 35ms before γ_{QRS-} and ending at γ_{QRS-}. The absolute value of that segment was computed, its median was removed, and the PT was applied with $R_v = 0.005$ in order to minimise the effect of interfering noise. Finally, the Q-wave was detected on that signal. To do so, the lowest local minimum in $\varphi[n]$ was found (in the case that $\varphi[n]$ had no minima, the lowest value was picked instead). Next, points in the window were identified which were both earlier than this minimum and had values of greater than 50% of the maximum value of $\varphi[n]$ within the window. If any points meeting these criteria were identified, then among the phasors exceeding the threshold, the one with the highest magnitude $M[n]$ was marked as the Q-wave. Otherwise, the Q-wave was identified as the picked minimum.

5.2.2.3 P-wave detection and delineation

Two approaches were implemented for P-wave detection and delineation: the PT and a wavelet-based algorithm [103]. By assessing their performance on the QT database [122], we found that, while there was some improvement in performance using the wavedet algorithm, both approaches gave satisfactory results. However, given that the performance of the PT was much better in patients prone to AF, it was used during the analysis. More details on the performance assessment of both P-wave delineators can be found in Appendix A.

P-wave detection

In order to detect the P-wave, a seek window previous to the Q-wave was established. The width of this window was normally one quarter of the last RR interval (i.e. the time between the current R-peak and the previous one), but in cases where this interval was greater than 1.5 times the mean RR (e.g., when one or more R-peaks were not detected), the width was defined as half of the mean RR. This allowed to compensate for errors in R-peak detection, averting for the detection of waves other than the P-wave. Next, the median from that ECG segment was removed, and the PT was applied with $R_v = 0.003$. The maximum value of $\varphi[n]$ within that window was located and if it was a local maximum in $M[n]$, that point was annotated as the P-wave peak (Figure 5.2). If the maximum was not associated with a M[n] peak, the start or end of the window was shortened by 15ms because part of the preceding T-wave or next Q-wave could have

been included within the initial window. This process was repeated iteratively until the detection of a P-wave peak occurred or the window length was null.

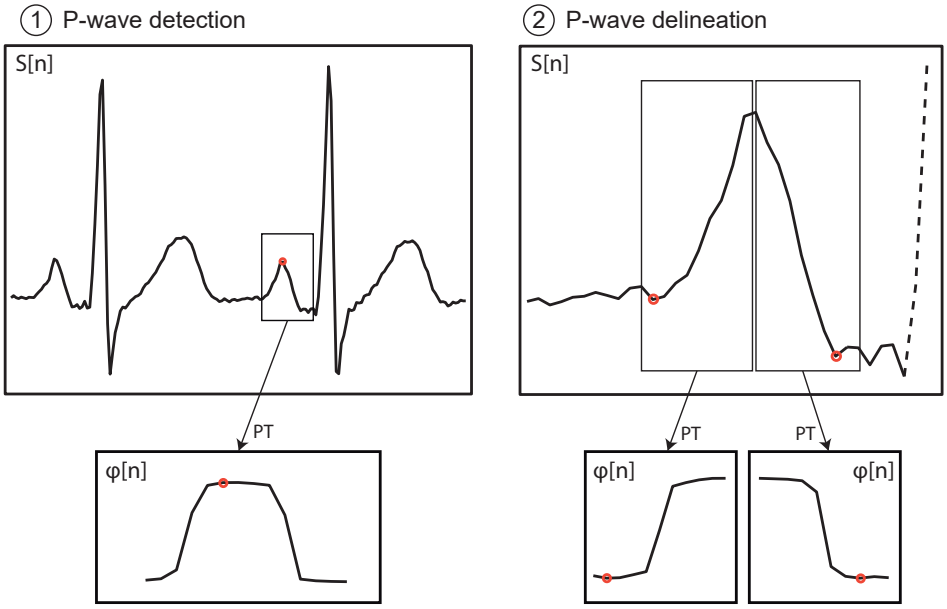

Figure 5.2: Representation of the P-wave detection and delineation process, using the phasor transform (PT). An adaptive window relative to the R-peak was used to detect the P-wave peak. Around this peak, two windows were established to delineate the P-wave onset and offset.

Finally, in contrast to the original algorithm where P-waves with amplitudes of less than 5% of the R-peak amplitude are discarded [104], all the detected P-waves were included at this point. Then, we used the PQI algorithm (described in Chapter 4) to exclude absent and low quality P-waves from the analysis. By excluding such P-waves, we ensured that the analysis was performed using only reliable measurements.

P-wave delineation

The detected P-wave peaks served as reference for the delineation of the P-wave. The onset and offset were thus individually searched. For the P-wave onset delineation, a 75ms window relative to the P-wave peak was established before that point. The median of that window was removed, and the PT was computed with $R_v = 0.005$ (Figure 5.2). Next, all the local minima were calculated. The closest minimum to the P-wave peak with at least 90% the value of the lowest amplitude minimum was found. If that point indeed existed and had an amplitude lower than 90% of the lowest value of the $\varphi[n]$ within that window, the P-wave onset was marked; otherwise, the process was repeated with a wider window of 115ms, as the onset was probably not contained within that initial window. This additional search procedure allowed for wider P-waves - an important improvement to the original algorithm. If the P-wave onset was still not found within that larger window, the two aforementioned search procedures were repeated with lower

values fo R_v, until the location of the onset was found, or while $R_v > 0$. Naturally, the P-wave offset delineation was based on the same process, with the difference that the search windows were established after the P-wave peak.

The process for negative P-waves was similar, but used negative values for R_v. Given that there were only 3 records with negative P-waves, this adaptation was performed manually.

Out of the 61 records included in the analysis, 26.2% needed their P-wave detection and delineation to be manually improved: 1.7% had the P-wave peak search window extended, and 24.5% had the P-wave onset/offset search window extended or shortened. Out of these records, 75.0% were from PAF patients.

5.2.3 Feature extraction

A selection of previously proposed ECG features were extracted in a beat-to-beat fashion (Table 5.1). The majority of the features were related to the P-wave, and have been previously used to predict AF. A time-series of feature measurements was extracted for each feature and for each patient (one data point per beat). Below we describe the more complex features.

5.2.3.1 Correlation coefficient index (CCI)

P-wave morphology variability was estimated by comparing each individual P-wave to a template [27]. The lower the correlation between the template and a individual P-wave, the more different they are, and vice-versa (Figure 5.3).

The methodology used to calculate this feature was adapted to be similar to that for the PQI. Briefly, the process started by extracting the P-wave signal between the corresponding onset and offset. Next, each P-wave was aligned and a template was created as in the PQI tool (refer to Subsection 4.2.2 for further detail). Finally, the correlation was computed using the cross-correlation with no temporal shift (i.e. point multiplication) between each P-wave, P_i, and the template, T:

$$CCI_i = \sum_{k=1}^{N} P_i[k]T[k] \tag{5.4}$$

5.2.3.2 Warping index (WI)

Dynamic time warping (DTW) was used to compute P-wave variability [27]. According to the DTW algorithm, the best alignment between two signals that are similar in shape but out of phase can be obtained by compressing or extending the time axis in a non-linear fashion. P-wave morphological variability was assessed by calculating the "amount" of modification required to maximise similarity between P-waves (Figure 5.3).

Given two P-waves, $X = \{x_1, x_2, ..., x_N\}$ and $Y = \{y_1, y_2, ..., y_M\}$, a matrix $G(N \times M)$ containing the distance between each point from X and each point from Y was built. Hence,

Table 5.1: Selection of features extracted from the electrocardiogram signal.

Feature	Description
Time features	
P_{dur}/$P_{dur,RRnorm}$	P-wave duration (either normalised with the current RR interval or not) [24, 25, 106, 123]
$P_{init.\ dur}$/$P_{fin.\ dur}$	P-wave initial/final duration: from onset to peak/from peak to offset, respectively [24, 25, 106]
PQ_{on}/ PQ_{off}/ $PQ_{on.RRnorm}$/ $PQ_{off.RRnorm}$	PQ interval measured from P-wave onset/offset, respectively (either normalised with current RR or not) [24, 123]
PR_{on}/ PR_{peak}/ PR_{off}/ $PR_{on,RRnorm}$/ $PR_{peak,RRnorm}$/ $PR_{off,RRnorm}$	PR interval measured from P-wave onset/peak/offset, respectively (either normalised with current RR or not) [24, 25, 123]
RR	Time interval between R-peaks
Morphology features	
P_{amp}/ $P_{amp,\ Rnorm}$	P-wave amplitude (either normalised with R-peak amplitude or not) [24, 26]
$P_{off\ amp}$	P-wave offset amplitude [24]
P_{magn}	P-wave magnitude: $\max\{P\} - \min\{P\}$ [106]
P_{assy}	P-wave assymetry: $P_{init.\ dur}/P_{fin.\ dur}$ [25]
P_{skew}	P-wave assymetry measured with skewness
$P_{min.vel}$/$P_{max.vel}$	P-wave minimum/maximum velocity [26]
$P_{vel.\ disp.}$	P-wave dispersion: $P_{min.vel} - P_{max.vel}$ [26]
P_{al}	P-wave arc-length [26]
$P_{rms,norm}$	P-wave root mean square (RMS), normalised with the number of samples [26]
P_{area}/$P_{area,norm}$	P-wave area (either normalised with P_{al} or not) [26]
P_{energy}/ $P_{energy,norm}$	P-wave energy (either normalised with P_{al} or not) [26]
$P_{energy\ ratio}$	P-wave energy ratio [26]
$P_{gauss.\ W}$/ $P_{gauss.\ C}$/ $P_{gauss.\ A}$/ $P_{gauss.\ error}$	P-wave Gaussian modelling parameters and error [105] (see Subsection 5.2.3.3 for details)
$P_{eucl.\ dist.}$	Euclidean distance between subsequent P-waves [124] (see Subsection 5.2.3.4 for details).
CCI	Correlation coefficient index [27] (see Subsection 5.2.3.1 for details)
WI_p, WI_t	Warping index [27] (see Subsection 5.2.3.2 for details)
PQ_{level}/ $PQ_{level,Rnorm}$/ $PQ_{level,Pnorm}$	PQ segment mean amplitude (either non-normalised or normalised with R-peak or P-wave peak amplitude) [24]
Q_{level}	Q-wave amplitude [24]

element $g[i,j]$ represented the relative distance between the two P-wave points x_i and y_j. The warping function (WF) was obtained, for a given cost function, by searching in matrix G the minimum cost path leading from point $g(N,M)$ to $g(1,1)$ [125, 126]. Finally, the warping index (WI) was the path length (expressed in number of samples), and was used to estimate variability [27]. The higher the WI, the more different are the two P-waves, and vice-versa.

The WI was extracted for every pair of consecutive P-waves (WI_p) and, additionally, between every P-wave and a P-wave template (created as in the CCI feature; WI_t).

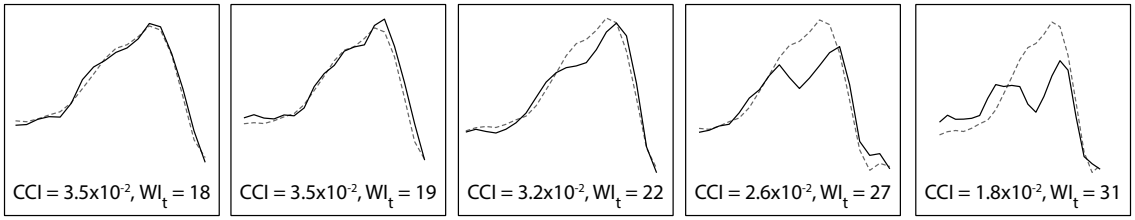

Figure 5.3: Examples displaying P-waves (solid line) and their template (dashed line), and correspondent correlation coefficient index (CCI) and warping index (WI_t). As indicated, increasing differences in morphology between P-waves and their template are associated with a decrease in CCI, and a increase in WI_t.

5.2.3.3 P-wave Gaussian modelling ($P_{gauss.}$)

In order to quantify morphology changes, each individual P-wave was modelled by a Gaussian function [105]. Hence, by comparing the Gaussian fitting parameters, we could assess several morphology features.

P-waves were modelled with a Gaussian function of the form:

$$\hat{p}[n] = A.e^{-((n-C)/W)^2} \tag{5.5}$$

where the constants A, C, and W represented amplitude, peak time-position, and width, respectively. Those parameters were computed by fitting the Gaussian function $\hat{p}[n]$ to the P-wave $p[n]$ by a non-linear least-squares approach.

Considering a logarithmic transformation, the P-wave signal $p[n]$ and its Gaussian model $\hat{p}[n]$ were related as:

$$\begin{aligned} ln(p[n]) &= ln(\hat{p}[n]) \\ &= \ln(A) - \left(\frac{n-C}{W}\right)^2 \end{aligned} \tag{5.6}$$

Defining the variables $b = ln(p[n])$, $a_2 = -1/W^2$, $a_1 = 2C/W^2$, and $a_0 = ln(A) - (C/W)^2$, Equation 5.6 can be written as a second-order polynomial:

$$\begin{aligned} b &= -\frac{1}{W^2}n^2 + \frac{2C}{W^2}n + ln(A) - \left(\frac{C}{W}\right)^2 \\ &= a_2 n^2 + a_1 n + a_0 \end{aligned} \tag{5.7}$$

Considering discrete values of n, coefficients a_2, a_1, and a_0 can be obtained by solving the following system of linear equation:

$$\begin{aligned} b(n_1) &= a_2 {n_1}^2 + a_1 n_1 + a_0 \\ b(n_2) &= a_2 {n_2}^2 + a_1 n_2 + a_0 \\ &\vdots \\ b(n_L) &= a_2 {n_L}^2 + a_1 n_L + a_0, \end{aligned} \tag{5.8}$$

where n_i is the i^{th} time instant for which the P-wave is defined. This system can be expressed in matrix notation as:

$$\mathbf{b} = \mathbf{N} \cdot \mathbf{a} \tag{5.9}$$

where $\mathbf{b} = \left[b[n_1], b[n_2], ..., b[n_L]\right]^T$, $\mathbf{a} = [a_2, a_1, a_0]^T$, and

$$\mathbf{N} = \begin{bmatrix} n_1^2 & n_1 & 1 \\ n_2^2 & n_2 & 1 \\ \vdots & \vdots & \vdots \\ n_L^2 & n_L & 1 \end{bmatrix} \tag{5.10}$$

By premultiplying Equation 5.9 by the matrix transpose $\mathbf{N}^T$, the vector $\mathbf{a}$ can be computed as:

$$\mathbf{a} = (\mathbf{N}^T.\mathbf{N})^{-1}.\mathbf{N}^T.\mathbf{b} \tag{5.11}$$

Once $\mathbf{a}$ was calculated, containing the indices a_2, a_1, and a_0, the constants A, C, and W could be calculated as follows:

$$W = \sqrt{-\frac{1}{a_2}} \tag{5.12}$$

$$C = \frac{a_1 W^2}{2} \tag{5.13}$$

$$A = e^{(a_0 + (C/W)^2)} \tag{5.14}$$

Finally, differences between each P-wave and its Gaussian model were quantified using the normalised root mean square error (Figure):

$$\epsilon = \sqrt{\frac{\sum_{k=1}^{L} \left|\hat{p}[k] - p[k]\right|^2}{\sum_{k=1}^{L} \hat{p}[k]^2}} \tag{5.15}$$

5.2.3.4 Euclidean distance between P-waves ($\mathbf{P_{eucl.\ dist.}}$)

The beat-to-beat euclidean distance between subsequent P-waves was used as an index of P-wave stability [124]. The greater the distance between P-waves, the more different they are, and vice-versa.

Firstly, the signal of each P-wave was extracted between the corresponding onset and offset, and the time-position of their centroid was calculated. Next, the P-waves were

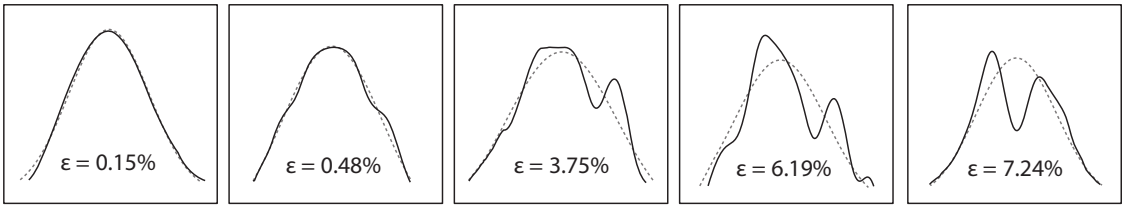

Figure 5.4: Examples displaying original P-waves (solid line) and their Gaussian modelling (dashed line). As can be appreciated, increasing differences in morphology are associated with increased root mean square error, ϵ. Adapted from Martinez *et al.* [105].

realigned with regard to their centroid (time-wise) and onset voltage (voltage-wise). The delineated P-waves had varying durations. For the purposes of calculating the Euclidean distance between P-waves we ensured that the first and second half of each P-wave had durations equal to the longest durations across the pair of P-waves. This was achieved by extracting longer periods of signal for the shorter P-waves in order to extract the desired P-wave duration. Finally, the euclidean distance between the two resynchronised P-waves was computed as follows:

$$P_{\text{eucl. dist., i}} = \frac{\sqrt{\sum_{k=1}^{N}(P_{i+1}[k] - P_i[k])^2}}{\sqrt{\sum_{k=1}^{N}(P_{i+1}[k])^2}}, \tag{5.16}$$

where $P_i[k]$ was the k^{th} sample of the i^{th} P-wave.

5.2.4 Metric calculation

The metric calculation stage allowed, for each method, to summarise each feature time-series as a single parameter. The obtained metrics were then used in the statistical analysis and in the prediction of AF. Implemented methods included the simple mean and standard deviation (SD), and several variability methods (Table 5.2), which will be described in depth below.

5.2.4.1 Linear variability time course

The time course of the features' variability on the 30-minute recordings was studied based on a linear regression method [25, 26, 105]. This method was used to understand how the features' variability progressed with time (Figure 5.5).

First, the time-series of a given P-wave feature was divided into segments of S samples. Next, a variability series was obtained by computing the difference between the 10^{th} and 90^{th} percentiles of each segment of S samples. In this way, outliers due to artefacts could be attenuated or rejected. An optimal value of S allowed slight variations in P-wave dynamics to be tracked, while minimising the effect of noise.

Table 5.2: Implemented methods, and corresponding computed metrics.

Method name	Metrics	Method explanation
Simple statistical metrics	• Mean • Standard deviation (SD)	Computes the mean value and the amount of variation (SD) of a feature's time-series.
Linear variability	• Slope of regression (α) • Median variability (m)	Calculates the variability of a feature's time-series, and studies how that variability progresses with time (increases [$\alpha > 0$], decreases [$\alpha < 0$] or maintains [$\alpha = 0$]).
Non-linear variability	• Central tendency measurement (CTM)	Computes the feature's time-series variability using difference plots. The higher the CTM, the lower the variability, and vice-versa.
P-wave amplitude dispersion	• Amplitude dispersion index (ADI)	Quantitative indicator of P-wave morphology variability. The greater the ADI, the greater the variability, and vice-versa. Uses the P-waves' signal.
Heart rate variability	• $\overline{RR}$, SDRR, SDSD, RMSSD, NN20, pNN20, NN50, pNN50 • TP, VLF, LF, HF, pHF, pLF, LF/HF	Studies the variability of the RR series using metrics from time- and frequency- domain. Gives indications on the autonomic nervous system.

$\overline{RR}$, mean RR; SDRR, standard deviation (SD) of RR; SDSD, SD of successive RR intervals; RMSSD, root mean square of differences between successive RR intervals; NN20/ NN50, number of RR pairs differing by at least 20/50ms, respectively; pNN20/ pNN50, percentage of NN20/ NN50, respectively; TP, total power; VLF, very low frequency power; LF, low frequency power; HF, high frequency power; pHF/ pHF, percentage of HF/ LF, respectively; LF/HF, ratio of LF to HF.

Finally, the features' variability trend was estimated by fitting a straight line to the variability series:

$$\hat{\mathbf{y}} = \alpha \mathbf{x} + \beta \tag{5.17}$$

where $x = \{1, 2, ..., N\}$, with N being the number of segments with S samples under analysis. The slope α was computed so that it minimised the linear regression model's squared residuals, i.e.:

$$\alpha = \frac{\sum_{i=1}^{N} (\boldsymbol{x}(i) - \bar{\boldsymbol{x}})(\boldsymbol{y}(i) - \bar{\boldsymbol{y}})}{\sum_{i=1}^{N} (\boldsymbol{x}(i) - \bar{\boldsymbol{x}})^2} \tag{5.18}$$

where $\bar{\boldsymbol{y}}$ and $\bar{\boldsymbol{x}}$ were the mean values of $\boldsymbol{y}$ and $\boldsymbol{x}$, respectively. The fitting slope α was therefore indicative of the feature variability time course: a positive or negative slope indicated increasing or decreasing variability, respectively, while a null slope indicated constant variability over time.

Finally, in addition to studying the variability time course, we have quantified the variability itself by taking the median value of the variability series (Figure 5.5).

The aforementioned slope and median variability metrics were computed for several values of S, which was varied from 10 to 40.

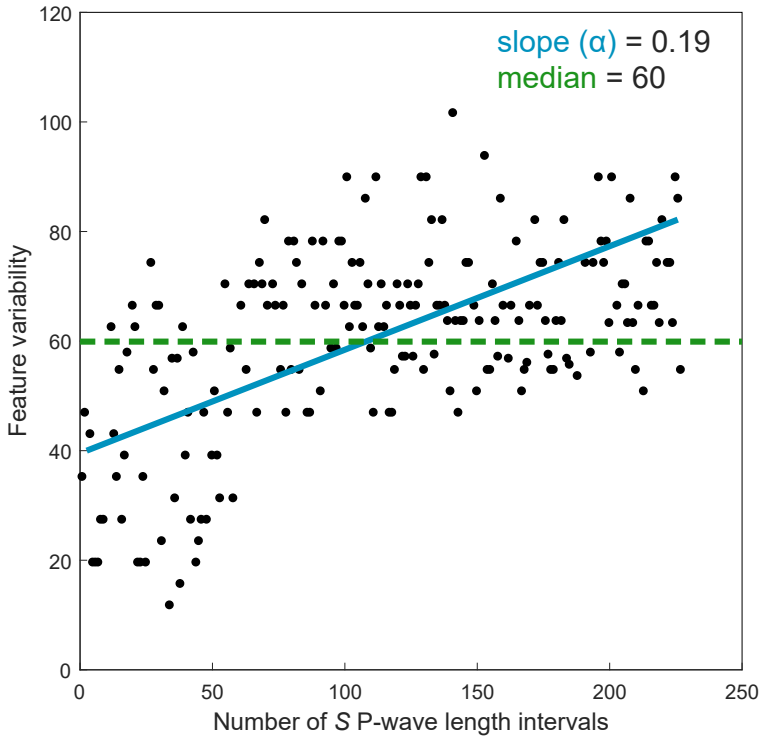

Figure 5.5: Two different metrics were extracted from the variability series: slope of linear fitting (blue) and median (green).

5.2.4.2 Non-linear variability

Alterations of the P-wave features' non-linear dynamics were quantified using the Central Tendency Measurement (CTM), a method which makes use of a difference-plot to compute variability [107, 127].

Given a feature time-series f and a lag d, the d^{th}-order difference plot corresponded to the plot $(f[n+d+1]-f[n+d])$ *vs.* $(f[n+1]-f[n])$: a dispersion plot centred around the origin (Figure 5.6). The CTM metric was computed as the proportion of points that lied within a circular region centred on the origin with radius ρ. Thus, a low CTM value reflected high variability, while a high CTM value reflected low variability.

From a mathematical point of view, given a feature time-series f with N data points, the dth order difference plot contained $N-d-1$ points. Hence, the d^{th}-lagged CTM was computed as:

$$CTM[d] = \frac{\sum_{i=1}^{N-d-1} \delta_d[i]}{N-d-1} \tag{5.19}$$

where

$$\delta_d[i] = \begin{cases} 1, & \text{if } D_d[i] < \rho \\ 0, & \text{otherwise} \end{cases} \tag{5.20}$$

with $D_d[i]$ being the euclidean distance from the i^{th} point from the d^{th}-order difference plot to the origin:

$$D_d[i] = \sqrt{(f[i+d+1] - f[i+d])^2 + (f[i+1] - f[i])^2} \tag{5.21}$$

Variability was studied by computing the CTM metric from lags $d = 1, 2, ..., 10$, and by varying the selection radius $\rho = 5.1, 5.2, ..., 10$ multiples of the standard deviation of the analysed data-series (f). Given that no guidelines exist to optimise ρ, normalisation with regard to the standard deviation provided translation and scale invariance [107].

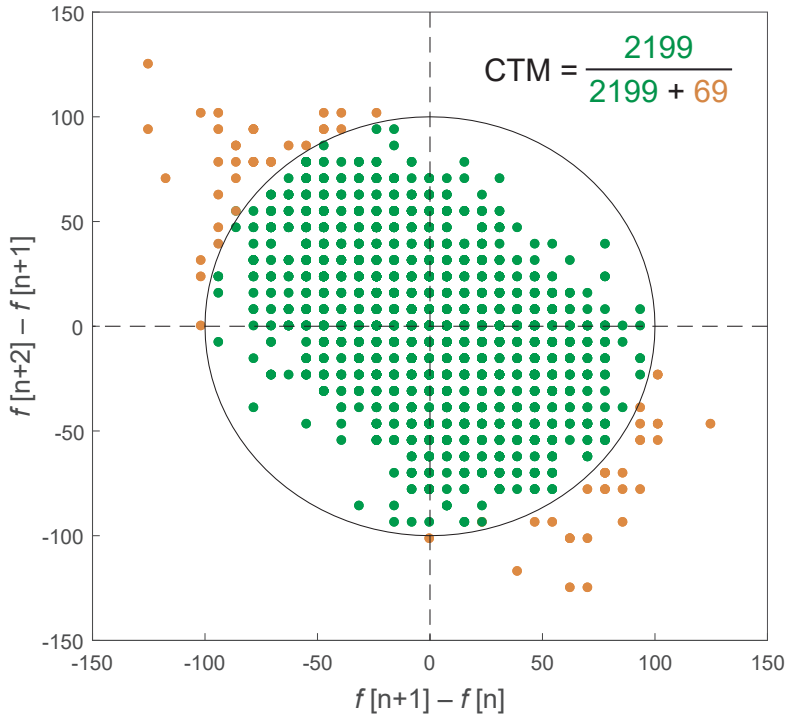

Figure 5.6: The central tendency measurement metric was calculated as the proportion of points that lied within a circular (green) region on the difference plot.

5.2.4.3 P-wave amplitude dispersion

The amplitude dispersion index (ADI) was used as a quantitative indicator of P-wave morphology variability [27]. The greater the ADI, the greater the variability, and vice-versa.

For each sample, the amplitude dispersion (AD) was calculated as the difference between the maximum and minimum amplitude values across all the P-waves (Figure 5.7):

$$AD[n] = max\{P[n]\} - min\{P[n]\} \tag{5.22}$$

where $max\{P[n]\}$ and $min\{P[n]\}$ corresponded to the maximum and minimum value of the n^{th} sample across all P-waves, respectively. Finally, the extracted index of P-wave

variability (ADI) was the maximum value of AD divided by the maximum amplitude value across all P-waves:

$$ADI = \frac{max\{AD\}}{max\{P\}} \tag{5.23}$$

Differently from the original study [27], where the ADI was calculated and averaged over the 12 ECG leads, we calculated the metric for a single ECG channel.

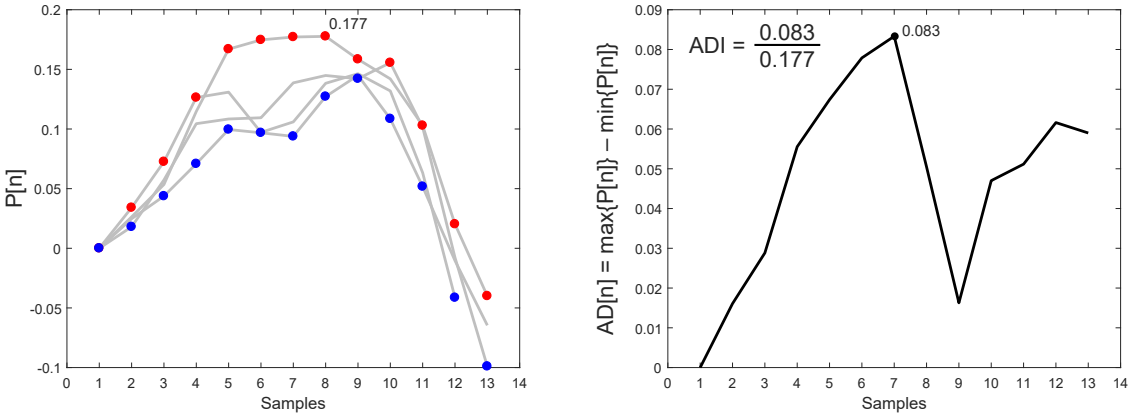

Figure 5.7: Calculation of the amplitude dispersion index (ADI) metric was performed in two steps. Firstly, the AD signal was computed by subtracting the maximum value across all P-waves (red) from the minimum values across all P-waves (blue). Finally, the ADI was calculated by dividing the maximum value of the AD signal by the maximum value across all P-waves.

5.2.4.4 Heart Rate Variability

Heart rate variability was studied in order to assess the role of the autonomic nervous system in the initiation of PAF. In order to eliminate artefacts in the RR-series (e.g., due to noise in the ECG signal, wrong R-peak detections, or atrial premature beats), a pre-processing filter was applied. This filter eliminated unreasonable RR intervals which differed by more than 20% from the mean of the current and previous RR interval, *rr*:

$$rr_i = \frac{2(RR_i - RR_{i-1})}{RR_i + RR_{i-1}} \tag{5.24}$$

Using the filtered RR series for the entire 30-minutes, several metrics from the time domain were calculated:

1. Time-domain metrics:

   a) $\overline{RR}$: Mean of the RR series;

   b) SDRR: Standard deviation of the RR series;

   c) SDSD: Standard deviation of successive differences between adjacent RR intervals;

d) RMSSD: Root mean square of differences between adjacent RR intervals:

$$RMSSD = \sqrt{\left(\frac{1}{N-1}\right)\sum_{i=1}^{N-1}(RR_{i+1} - RR_i)^2} \tag{5.25}$$

e) NN20/pNN20: Number and percentage of RR pairs differing by at least 20ms;

f) NN50/pNN50: Number and percentage of RR pairs differing by at least 50ms.

In addition, after resampling of the given RR-series using spline interpolation, several frequency-metrics were extracted using the fast Fourier transform (FFT):

2. Frequency domain metrics:

   a) TP: Total power between the frequencies 0 and 0.4Hz;

   b) VLF: Very low frequency power, calculated between the frequencies 0 and 0.04Hz;

   c) LF: Low frequency power, between the frequencies 0.04 and 0.15Hz;

   d) HF: High frequency power, between the frequencies 0.15 and 0.4Hz;

   e) pHF: percentage of power on the high frequency band;

   f) pLF: percentage of power on the low frequency band;

   g) LF/HF: low frequency power to high frequency power ratio;

5.2.5 Statistical analysis

Two comparisons were made: healthy controls *vs.* PAF patients, and PAF far *vs.* PAF close. Non-parametric Wilcoxon rank sum tests (Mann-Whitney U-test) were performed, and metrics with a statistical significance of $p < 0.01$ were considered as significant. Finally, a Holm-Sidak correction [128, 129] was performed at a significance level of 0.05. For the methods where several configurations were tested (linear and non-linear variability), this statistical correction was performed for the metrics within each variation (e.g., for all the linear variability metrics using $S = 10$). The presented significance levels were selected in order to minimise the probability of type I errors (false rejection of a null hypothesis). The uncorrected $p < 0.01$ allowed to find trends, while the corrected $p < 0.05$ allowed to find results with more concrete evidence of statistical significance (i.e., results even more likely to be true positives), as is commonly done in neuroimaging research [130].

For the methods where several variations were tested, the optimal variation was chosen as the one providing the highest number of metrics significant after correcting for multiple comparisons at $p < 0.05$. In case of a tie, the optimal variation was the one providing the highest number of significant metrics at uncorrected $p < 0.01$. Once the optimal variation (optimal S for the linear variability, and lag for the non-linear variability) was chosen for each comparison, it was used in the remaining analysis.

5.2.6 Performance assessment

The ability of each single statistically significant metric to discriminate between the different groups was assessed using the area under the receiver-operator curve (AUC) statistic, evaluated through comparisons of the predicted records' label and the corresponding true categories.

Finally, two different decision trees were assembled, one for each group comparison. These models were built in order investigate the non-monotonic relationships among single parameters, thus aiming to improve group classification. All the features which presented statistically significant differences were considered as candidate features for the classification trees. The stopping criterion used for the three growth was that each node contained fewer than 5% of all the observations. Performance was assessed using 10-fold cross validation, and was evaluated using the accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) statistics.

5.3 Results

5.3.1 Simple statistical metrics

Table 5.3 shows the significant (uncorrected $p < 0.01$) results obtained when comparing healthy controls to PAF patients using the simple mean and standard deviation metrics. The mean values of $PQ_{on.RRnorm}$, PR_{on}, $PR_{on\ RRnorm}$, PR_{peak}, $PR_{peak.RRnorm}$, $PR_{off.RRnorm}$, P_{amp}, and $P_{amp,\ Rnorm}$ were found to be significantly greater in PAF patients than in controls. The same was found for the standard deviation of PR_{off}, $PR_{off.RRnorm}$, and $PR_{on.RRnorm}$. No result was statistically significant after correcting for multiple comparisons. Furthermore, no significant results were found when comparing PAF far and PAF close.

5.3.2 Linear Variability

The optimal number of samples (S) used to compute variability was optimized for each of the subject comparisons: $S = 35$ for healthy *vs.* PAF patients, and $S = 10$ for PAF far *vs.* PAF close.

PAF patients had significantly higher median variability than controls in $P_{amp,\ Rnorm}$ while, in opposition, healthy subjects had higher median variability in RR (Table 5.4). PAF close was found to have increasing variability, while the reverse was found for PAF far in P_{dur}, PQ_{off}, $PQ_{off\ RRnorm}$, PR_{off}, $PR_{off\ RRnorm}$, P_{amp}, P_{al}, P_{area}, CCI, WI_t, and $P_{eucl.\ dist.}$ (Figure 5.8). No results remained significant after correction for multiple comparisons.

5.3.3 Non-linear Variability

The 1st order (lag = 1) difference maps were used to compute the CTM metrics for both subject comparisons. Table 5.5 shows the metrics reaching significance, accompanied

Table 5.3: Significant results (uncorrected $p < 0.01$) obtained when testing the simple statistical metrics and the heart rate variability analysis. Results significant after correction for multiple corrections (corrected $p < 0.05$) are shaded in grey. The median value (and the first/ third quartiles) of each group are presented.

Healthy *vs.* PAF patients				
Metrics	**Healthy**	**PAF**	***P*-value**	**AUC**
$\overline{\mathbf{PQ_{on.RRnorm}}}$ [a.u.]	0.172 (0.149 / 0.191)	0.209 (0.187 / 0.265)	4.9×10^{-4}	0.76
$\overline{\mathbf{PR_{on}}}$ [ms]	209.175 (187.982 / 225.827)	226.356 (222.170 / 244.444)	9.5×10^{-3}	0.69
$\overline{\mathbf{PR_{on\ RRnorm}}}$ [a.u.]	0.256 (0.227 / 0.283)	0.297 (0.280 / 0.368)	3.7×10^{-4}	0.77
$\overline{\mathbf{PR_{peak}}}$ [ms]	150.824 (131.058 / 164.620)	180.936 (167.000 / 188.701)	2.1×10^{-3}	0.73
$\overline{\mathbf{PR_{peak.RRnorm}}}$ [a.u.]	0.184 (0.151 / 0.202)	0.227 (0.197 / 0.275)	1.4×10^{-4}	0.78
$\overline{\mathbf{PR_{off.RRnorm}}}$ [a.u.]	0.114 (0.090 / 0.139)	0.146 (0.126 / 0.171)	1.9×10^{-3}	0.73
$\overline{\mathbf{P_{amp}}}$ [mV]	0.078 (0.033 / 0.159)	0.151 (0.094 / 0.201)	2.5×10^{-3}	0.73
$\overline{\mathbf{P_{amp,\ Rnorm}}}$ [a.u.]	0.045 (0.014 / 0.072)	0.085 (0.069 / 0.123)	1.1×10^{-3}	0.74
$\mathbf{SD(PR_{off})}$ [ms]	10.570 (7.627 / 13.151)	16.199 (8.690 / 18.541)	8.4×10^{-3}	0.70
$\mathbf{SD(PR_{off.RRnorm})}$ [a.u.]	0.018 (0.014 / 0.022)	0.025 (0.017 / 0.029)	4.0×10^{-3}	0.72
$\mathbf{SD(PR_{on.RRnorm})}$ [a.u.]	0.027 (0.023 / 0.034)	0.021 (0.019 / 0.026)	6.5×10^{-3}	0.70
SDRR [ms]	0.085 (0.049 / 0.123)	0.048 (0.025 / 0.060)	2.2×10^{-4}	0.78
LF [ms^2]	0.718 (0.600 / 1.247)	0.565 (0.527 / 0.682)	1.2×10^{-3}	0.74

SD, standard deviation.

Table 5.4: Significant results (uncorrected $p < 0.01$) obtained when using the linear variability metrics. No results remained significant after correction for multiple comparisons. The median value (and the first/ third quartiles) of each group are presented.

Healthy *vs.* PAF patients ($S = 35$)				
Metrics	**Healthy**	**PAF**	***P*-values**	**AUC**
$\boldsymbol{m}\mathbf{(RR)}$ [ms]	99.6 (62.5 / 824.2)	46.9 (29.3 / 206.4)	3.3×10^{-3}	0.72
$\boldsymbol{m}\mathbf{(P_{amp,\ Rnorm})}$ [a.u.]	2.9×10^{-2} (2.3×10^{-2} / 4.2×10^{-2})	4.2×10^{-2} (3.1×10^{-2} / 6.0×10^{-2})	4.8×10^{-3}	0.71
PAF far *vs.* PAF close ($S = 10$)				
Metrics	**PAF far**	**PAF close**	***P*-values**	**AUC**
$\boldsymbol{\alpha}\mathbf{(P_{dur})}$ [ms]	-1.6×10^{-2} (-3.8×10^{-2} / -5.7×10^{-3})	9.3×10^{-3} (1.5×10^{-3} / 3.6×10^{-2})	9.4×10^{-3}	0.79
$\boldsymbol{\alpha}\mathbf{(PQ_{off})}$ [ms]	-2.2×10^{-2} (-3.7×10^{-2} / 6.4×10^{-3})	1.7×10^{-2} (3.7×10^{-3} / 6.6×10^{-2})	8.3×10^{-3}	0.79
$\boldsymbol{\alpha}\mathbf{(PQ_{off\ RRnorm})}$ [a.u]	-2.7×10^{-5} (-5.2×10^{-5} / -1.3×10^{-5})	1.8×10^{-5} (-1.5×10^{-6} / 7.0×10^{-5})	5.6×10^{-3}	0.80
$\boldsymbol{\alpha}\mathbf{(PR_{off})}$ [ms]	-2.7×10^{-2} (-5.2×10^{-2} / -9.9×10^{-3})	1.4×10^{-2} (4.3×10^{-3} / 5.3×10^{-2})	4.9×10^{-3}	0.81
$\boldsymbol{\alpha}\mathbf{(PR_{off\ RRnorm})}$ [a.u]	-3.7×10^{-5} (-5.9×10^{-5} / -4.4×10^{-6})	2.2×10^{-5} (7.4×10^{-6} / 6.0×10^{-5})	6.4×10^{-3}	0.80
$\boldsymbol{\alpha}\mathbf{(P_{amp})}$ [mV]	-5.0×10^{-5} (-5.5×10^{-5} / 5.8×10^{-6})	8.1×10^{-5} (-2.2×10^{-5} / 1.8×10^{-4})	6.4×10^{-3}	0.80
$\boldsymbol{\alpha}\mathbf{(P_{al})}$ [μVs]	-2.0×10^{-3} (-4.8×10^{-3} / -7.4×10^{-4})	1.2×10^{-3} (1.9×10^{-4} / 4.6×10^{-3})	9.4×10^{-3}	0.79
$\boldsymbol{\alpha}\mathbf{(P_{area})}$ [μVs]	-1.8×10^{-4} (-8.8×10^{-4} / -4.6×10^{-5})	2.6×10^{-4} (6.8×10^{-5} / 5.2×10^{-4})	4.2×10^{-3}	0.81
$\boldsymbol{\alpha}\mathbf{(CCI)}$ [a.u]	-3.4×10^{-5} (-4.8×10^{-5} / -1.4×10^{-6})	2.1×10^{-5} (-2.1×10^{-6} / 5.6×10^{-5})	9.4×10^{-3}	0.79
$\boldsymbol{\alpha}\mathbf{(WI}_t\mathbf{)}$ [samples]	-3.5×10^{-3} (-9.1×10^{-3} / 4.0×10^{-4})	4.0×10^{-3} (6.5×10^{-4} / 6.1×10^{-3})	3.2×10^{-3}	0.82
$\boldsymbol{\alpha}\mathbf{(P_{eucl.\ dist})}$ [a.u]	-9.2×10^{-5} (-1.4×10^{-4} / -2.8×10^{-6})	8.4×10^{-5} (9.9×10^{-7} / 1.6×10^{-4})	3.2×10^{-3}	0.82

m, mean variability; α, slope of regression.

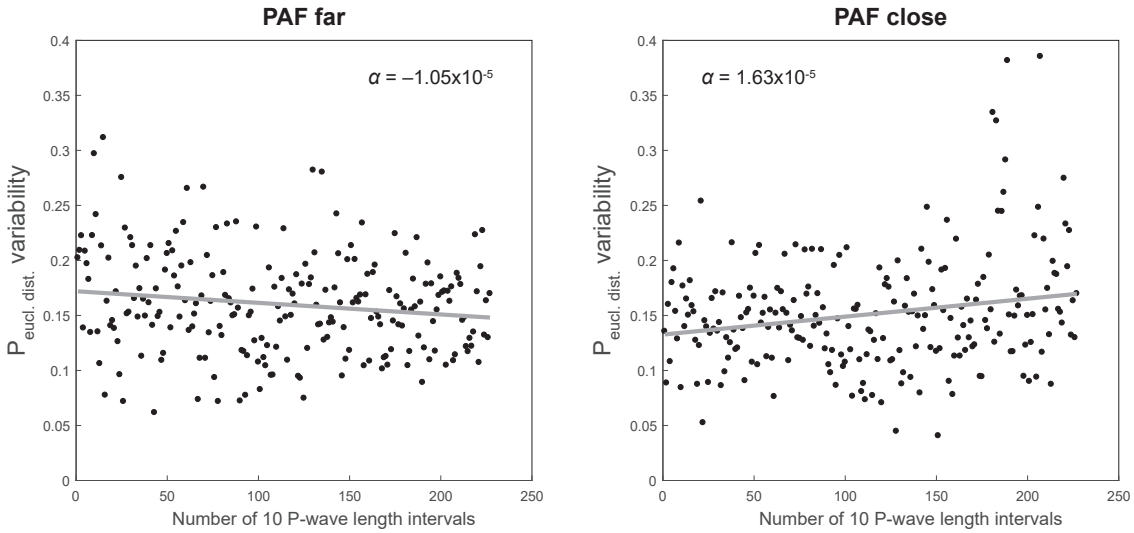

Figure 5.8: Illustrative example of P-wave euclidean distance variability time course from a typical patient far away from the onset of paroxysmal atrial fibrillation (PAF; left) and from a typical patient close to the arrhythmia onset (right). Patients close to the onset of PAF had a significantly higher variability slope than patients far away from the arrhythmia onset.

by the associated optimal radius, ρ. The CTM calculated over PQ_{off}, PQ_{on}, $PQ_{on.RRnorm}$, PR_{off}, $PR_{off.RRnorm}$, $PR_{on.RRnorm}$, $PR_{peak\ RRnorm}$, P_{amp}, $P_{amp,\ Rnorm}$, $P_{gauss.\ A}$, P_{magn}, $P_{min.vel}$, $P_{area,norm}$, and CCI was significantly higher in healthy individuals than in PAF. From those, only $PR_{on.RRnorm}$, P_{amp}, $P_{amp,\ Rnorm}$, P_{magn}, $P_{min.vel}$, and CCI remained significant after correction for multiple corrections. In addition, the CTM calculated over $P_{gauss.\ A}$ was significantly higher in PAF far than in PAF close. Finally, no CTM metric was significantly higher in PAF (*vs.* healthy) nor in PAF close (*vs.* PAF far).

5.3.4 P-wave amplitude dispersion

The ADI metric was not found to significantly differ between any of the groups.

5.3.5 Heart rate variability

Healthy subjects had significantly higher SDRR and LF than PAF patients (Table 5.3), which was still significant after correction for multiple comparisons. No HRV metric was significantly different between PAF far and PAF close.

5.3.6 Performance assessment

The classification ability of each individual significant metric, measured using the AUC, is presented in Tables 5.3 to 5.5. Finally, the decision trees built using the statistically significant metrics are shown in Figures 5.10 and 5.11, and their classification performance is presented in Table 5.6. Both decision trees had 0.88 accuracy, 0.82 sensitivity, and 0.93

Table 5.5: Significant results (uncorrected $p < 0.01$) obtained when testing the non-linear variability results. The presented values of optimal ρ correspond to the number of multiples of the standard deviation which minimised the p-value for the correspondent feature. Results significant after correction for multiple corrections (corrected $p < 0.05$) are shaded in grey. The median value (and the first/ third quartiles) of each group are presented (in arbitrary units).

Healthy *vs.* PAF patients (*lag* = 1)					
Features	**Optimal ρ**	**Healthy**	**PAF**	***P*-value**	**AUC**
$\mathbf{PQ_{off}}$	6.0	1.0000 (0.9998 / 1.0000)	0.9995 (0.9970 / 1.0000)	6.0×10^{-3}	0.68
$\mathbf{PQ_{on}}$	5.5	0.9966 (0.9933 / 0.9991)	0.9943 (0.9886 / 0.9962)	8.4×10^{-3}	0.70
$\mathbf{PQ_{on.RRnorm}}$	6.0	0.9995 (0.9980 / 1.0000)	0.9977 (0.9953 / 0.9987)	4.2×10^{-4}	0.76
$\mathbf{PR_{off}}$	10.5	1.0000 (1.0000 / 1.0000)	1.0000 (0.9988 / 1.0000)	8.7×10^{-3}	0.62
$\mathbf{PR_{off.RRnorm}}$	9.5	1.0000 (0.9994 / 1.0000)	0.9989 (0.9978 / 1.0000)	5.2×10^{-3}	0.69
$\mathbf{PR_{on.RRnorm}}$	5.0	0.9977 (0.9930 / 0.9998)	0.9874 (0.9787 / 0.9937)	2.0×10^{-4}	0.78
$\mathbf{PR_{peak\ RRnorm}}$	6.5	1.0000 (0.9974 / 1.0000)	0.9932 (0.9893 / 0.9993)	4.1×10^{-4}	0.75
$\mathbf{P_{dur.\ RRnorm}}$	5.5	0.9985 (0.9957 / 1.0000)	0.9949 (0.9905 / 0.9978)	1.7×10^{-3}	0.73
$\mathbf{P_{amp}}$	5.0	0.9936 (0.9903 / 0.9991)	0.9852 (0.9803 / 0.9922)	6.7×10^{-5}	0.80
$\mathbf{P_{amp,\ Rnorm}}$	5.0	0.9934 (0.9902 / 0.9984)	0.9870 (0.9836 / 0.9910)	1.1×10^{-4}	0.79
$\mathbf{P_{gauss.\ A}}$	5.0	0.9969 (0.9927 / 0.9985)	0.9905 (0.9873 / 0.9966)	4.1×10^{-3}	0.71
$\mathbf{P_{magn}}$	6.0	0.9993 (0.9955 / 1.0000)	0.9966 (0.9918 / 0.9985)	7.4×10^{-4}	0.75
$\mathbf{P_{min.vel}}$	7.5	1.0000 (1.0000 / 1.0000)	0.9981 (0.9971 / 0.9996)	5.8×10^{-5}	0.78
$\mathbf{P_{area,norm}}$	5.0	0.9947 (0.9894 / 0.9992)	0.9898 (0.9860 / 0.9925)	2.9×10^{-3}	0.72
CCI	5.0	0.9969 (0.9924 / 0.9995)	0.9893 (0.9843 / 0.9951)	6.4×10^{-4}	0.75
PAF far *vs.* PAF close (lag = 1)					
Metrics	**Optimal ρ**	**PAF far**	**PAF close**	***P*-value**	**AUC**
$\mathbf{P_{gauss.\ A}}$	6.0	0.9985 (0.9971 / 0.9990)	0.9933 (0.9888 / 0.9968)	1.3×10^{-3}	0.85

ρ, radius used to compute CTM.

specificity. These decision trees included variability metrics (regression slope and CTM) applied to several P-wave time- and morphological features.

Table 5.6: Classification results obtained when comparing healthy subjects to patients who developed paroxysmal atrial fibrillation (Healthy *vs.* PAF), and patients far and close to the arrhythmia onset (PAF far *vs.* PAF close).

Metric	**Healthy *vs.* PAF**	**PAF far *vs.* PAF close**
Accuracy	0.88	0.88
Sensitivity	0.82	0.82
Specificity	0.93	0.93
PPV	0.94	0.93
NPV	0.83	0.85

PPV, positive predictive value; NPV, negative predictive value.

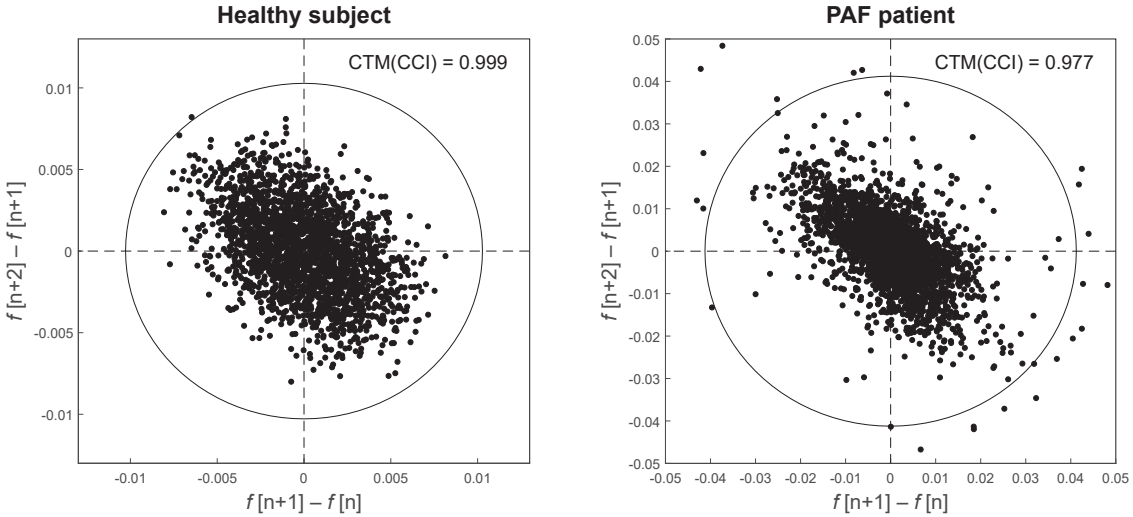

Figure 5.9: Illustrative example of difference plots from a typical healthy subject (left) and a patient who developed paroxysmal atrial fibrillation (PAF; right). Patients who developed PAF had significantly higher non-linear variability (as indicated using the central tendency measurement; CTM) than healthy subjects.

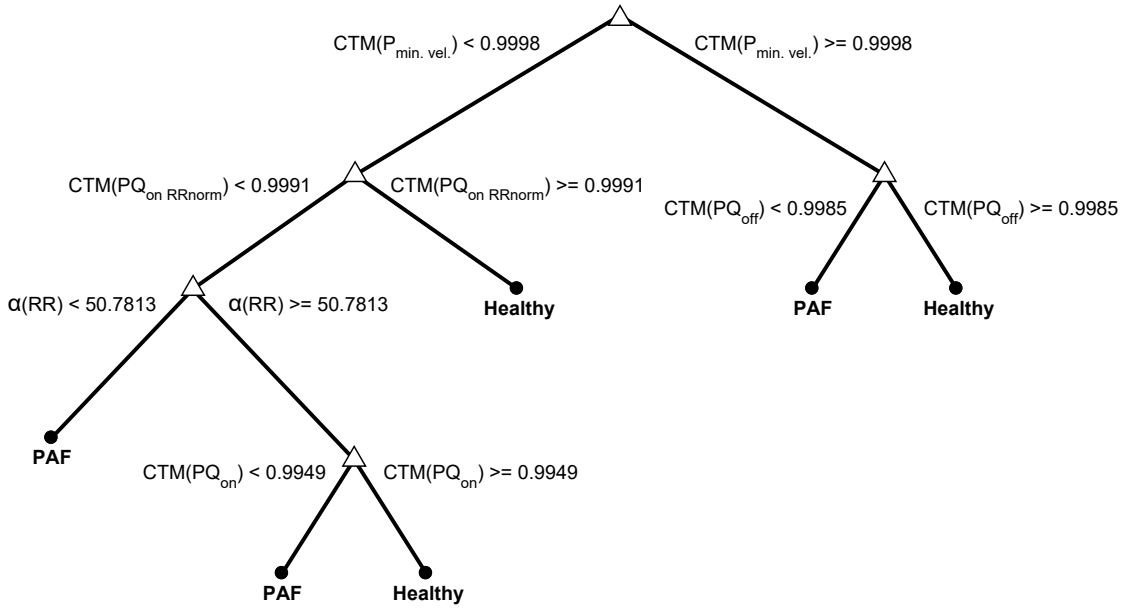

Figure 5.10: Decision tree built to compare healthy subjects (Healthy) to patients which developed paroxysmal atrial fibrillation (PAF).

5.4 Discussion

Prediction of PAF is a clinically important challenge, because the loss of sinus rhythm may by prevented using prophylactic treatments. Maintenance of sinus rhythm can lead to improved hemodynamics and decreased symptoms and complications. Furthermore, by avoiding the arrhythmia onset, there is a decrease in atrial anatomic and electrical remodelling that causes increased susceptibility in future episodes of PAF [131]. Hence,

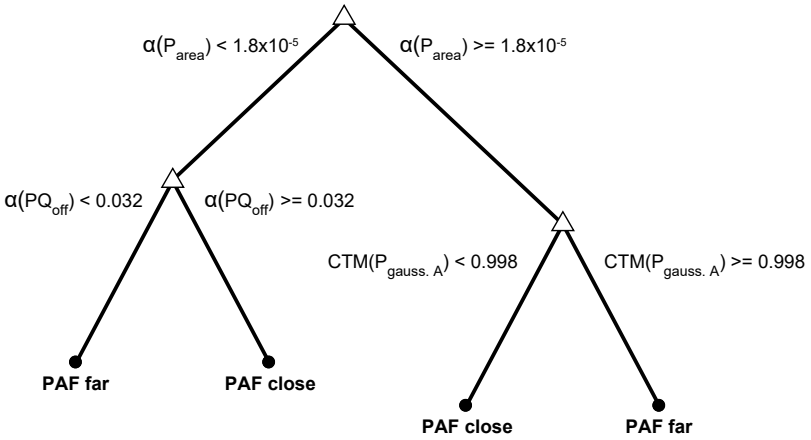

Figure 5.11: Decision tree built to distinguish between patients with paroxysmal atrial fibrillation that are far (PAF far) and close (PAF close) to the arrhythmia onset.

in the present chapter we aimed to identify individuals at risk of developing PAF based on their ECG, and to predict the imminent onset of AF in individuals known to be at risk, following the 2001 CinC challenge.

Several ECG features and metrics have been previously proposed to predict AF. However, a determination of which AF prediction methodology, if any, is more suitable for clinical use, has never been performed. Thus, in the present study we have implemented and refined previously proposed ECG features and metrics, and have tested if their combination improves the prediction of PAF. Finally, we ensured that only high quality P-waves were included in the analysis by using the novel automated PQI tool.

We have identified several P-wave and HRV metrics to be predictors of PAF. P-wave variability was found to be the best predictor, as it gave the highest number of significant metrics and had the best overall classification ability, measured using the AUC (Tables 5.4 and 5.5). Furthermore, the built decision trees mostly contained P-wave variability metrics (Figures 5.10 and 5.11). Like previous studies, we have found PAF to be associated with higher P-wave related variability [25–27, 105, 107–109, 124, 127]. In addition, that variability was significantly higher just before the arrhythmia onset [25, 26, 105, 107–109, 127]. Finally, the combination of variability metrics was able to identify subjects at risk of PAF and to predict the imminent onset of PAF with high specificity (0.93) and good sensitivity (0.82) and accuracy (0.88). These results are better than those from the majority of other beat-to-beat analyses (where accuracy ranged from 0.70 to 1.00; Table 3.3).

Unexpectedly, we have found the linear variability to decrease with time in PAF patients far away from the arrhythmia onset (Table 5.4, Figure 5.8). This effect was found across several P-wave time and morphology metrics. A possible explanation for this trend of decreased variability might be related to the timing in which the ECG signals from PAF far patients were selected. Even though the recordings were extracted at least 45 minutes after the previous PAF episode, it might be that the heart was still recovering

from the big impact that AF causes, which could explain the decrease of variability with time. However, this hypothesis cannot be confirmed due to the lack of clinical information contained within the database, and due to the lack of research into P-wave variability during the recovery from AF.

Furthermore, unexpectedly, we did not found significant differences between Healthy and PAF patients when using linear variability regression. This was probably due to the fact that the PAF group contained subjects both from PAF close (with increasing variability, $\alpha > 0$) and PAF far (with decreasing variability, $\alpha < 0$), which, together, led to no significant differences when compared with Healthy subjects, which we expected to have constant variability ($\alpha = 0$) [25]. Nonetheless, as expected, non-linear variability was augmented (i.e. lower CTM) in PAF (*vs.* Healthy) and in PAF close (*vs.* PAF far).

Finally, it is important to remark that, from a relatively large selection of features (Table 5.1), only a few were found to be significant. This might be due to the relatively low sampling frequency (128Hz), which might have hindered some subtle P-wave modifications that have been quantified at large sampling frequencies (1000Hz) in all previous studies. Hence, the P-wave features which we have found to be significant are probably the ones which are more pronounced and could be quantified at the present sampling frequency. Indeed, the lack of high resolution ECG devices has been identified as an obstacle for the wide clinical application of those proposed methods [27]. Hence, in the present study, we identified features which were capable of predicting AF in commonly used, low resolution ECG devices.

5.4.1 P-wave time analysis

Out of all the features, P-wave duration is accepted as the most reliable marker for atrial conduction characterization. However, even though P-waves with prolonged duration have been extensively associated with AF [23], other studies have failed to demonstrate any remarkable P-wave prolongation [132]. Taken together, these results suggest P-wave prolongation to be only a marker of atrial conduction slowing, and not an unavoidable requirement for AF development [105, 132]. Even though we did not found P-wave duration to be significantly longer prior to PAF, we found prolonged PQ and PR intervals in patients susceptible to PAF, suggesting decreased conduction velocity of atrial impulses until ventricular depolarization in that group.

Atrial depolarization is reflected in the ECG of patients prone to AF as the result of an interplay of two overlapping effects: 1) a decrease in conduction velocity of atrial impulses, causing prolongation of P-wave duration; and 2) overlap between atrial depolarization and possible premature atrial repolarization due to decreased refractory period, shortening P-wave duration. Hence, given the heterogeneous combination of both these effects on patients at risk of AF, greater variability in P-wave time intervals was expected in those patients. Indeed, we found P-wave duration and PQ and PR intervals to be associated with higher variability in PAF patients (*vs.* healthy subjects) and in PAF close (*vs.*

PAF far). These results are in line with previous studies showing the (non-linear and the slope of the linear) variability of several P-wave time features to be greater in patients with PAF, and within those patients, to increase as the arrhythmia onset approximated [25, 107].

5.4.2 P-wave morphology analysis

P-wave morphology reflects the atrial depolarization in a three-dimensional space. It is the result of a complex interplay of several factors such as [132]: 1) the origin of the sinus beat and right atrial depolarization; 2) left atrial depolarization; and 3) shape and size of the atrial chambers, which affect both the time and course of the depolarization process. Alterations in P-wave morphology can reflect disturbances in the atrial depolarization wavefront [133], which are, for instance, correlated with the induction of PAF [132].

Hence, our results of higher P-wave morphology variability are in agreement with the disturbed conduction observed in the atrial myocardium susceptible to AF [105]. Electrophysiological alterations within the atria, such as the slow propagation velocity of atrial impulses, inhomogeneous atrial activation and shortening of atrial refractory period, precede the onset of AF [24]. Moreover, structural abnormalities within the atrial myocardium, such as fibrosis, together with site-specific conduction delays, account for a nonuniform anisotropic propagation of sinus impulses. This inhomogeneous conduction is thought to play a major role in the initiation of reentry, due to the increased likelihood of unidirectional block of premature impulses [132]. Together, all these determinants invariably alter the way through which the sinus impulse travels across the atria, resulting in highly variable P-wave morphology [105].

Our results are in line with a large number of studies showing altered P-wave morphology to be an indicator of AF. Those studies have identified patients with fractioned conduction to be at risk of AF firstly by using signal-averaged P-waves [23], and later, like us, on a beat-to-beat basis. Indeed, by looking at each individual P-wave, it was possible to capture subtle changes in P-wave morphology. Firstly, the non-linear variability and the regression of variability of several P-wave morphological features, in both time [26] and frequency [108, 109] domains, were found to distinguish between healthy subjects, PAF far, and PAF close. Secondly, the metrics ADI, mean CCI and WI_p [27], and mean and SD of P-wave euclidean distance [124] were found to be predictors of PAF. In the present study, we have extended those features by calculating their variability which, indeed, were found to be significant, and to be better predictors than the simple original proposed metrics, which were not significant (e.g., the mean value of CCI was not significant, but its variability was). Finally, the identification of a distinct secondary P-wave morphology, together with the analysis of the symlet wavelet energy, enabled the identification of PAF patients in a recent study [120].

5.4.3 Heart rate variability

During normal sinus rhythm, the heart rate (HR) varies from beat to beat. HRV results from the dynamic interplay between multiple physiologic mechanisms that regulate the instantaneous HR. Since short-term HR regulation is predominantly governed by sympathetic and parasympathetic neural activity, study of HR fluctuations provides a way to observe the state and integrity of the autonomic nervous system [134]. Furthermore, autonomic nervous system activation can induce significant and heterogeneous changes of atrial electrophysiology and induce atrial tachyarrhythmias, such as AF [135]. Hence, we have studied the HRV in order to assess the role of the autonomic nervous system in PAF.

We found the SDRR to be decreased in PAF patients. This result, together with the decreased median variability of the RR series, indicate the impaired ability of that group to adapt to changing circumstances [134], which may predispose to untoward cardiac events [136]. Our results are in line with a previous finding of increased stability (decreased complexity) in PAF patients, using approximate entropy and detrended fluctuation analysis [136]. Furthermore, other studies have shown lower complexity in episodes preceding PAF compared to distant ones using several entropy indices [137, 138] and difference plots [138].

The control group also exhibited higher LF than PAF patients. This result is in opposition to previous findings of increased LF in patients who developed postoperative AF [139], and in PAF patients immediately before the arrhythmia (compared to distant ones) [137]. The LF component is directly associated with both the sympathetic and vagal activity, which, for instance, have been associated with atrial fibrillation [135, 140]. For instance, studies have revealed an important role of the sympathetic nervous system and its complex interaction with the vagal system in triggering AF [140]. Nonetheless, it should be pointed that the entire autonomic nervous system plays a role in the induction of AF. Finally, given that no clinical information was provided, we could not asses the influence of extraneous factors on this effect of increased LF in healthy subjects.

5.4.4 Limitations and Future Work

Even though the freely available CinC 2001 database stimulated research on the prediction of AF, it presents some disadvantages. Firstly, the ECG signals were not recorded with the special intent of studying the P-waves, which led us to reduce the sample size, and to analyse and compare P-wave metrics from different leads. Secondly, the effects of confounding factors, such as age and gender, could not be controlled, given that clinical data was not provided. Finally, signals were sampled at 128Hz, which might have hindered an accurate P-wave characterization in both time and morphological domain. Nonetheless, this sampling frequency is commonly used in hospital bedside monitors.

It is also important to note that some features, which were previously found to be predictive of AF, were not implemented. For instance, the P-waves' frequency [108, 109]

and time-frequency [120] contents could not be used to study their morphology, given that the sampling frequency with which the ECG signals were acquired was too low (a minimum of 400Hz are required to capture the frequency content up to 200Hz which was used in these studies, according to the Nyquist theorem). Furthermore, even though the number of ectopic beats has shown to be predictive of PAF [94], that method is only feasible just before the arrhythmia starts, which is too late for the application of an efficient prophylactic medication [136, 141].

Furthermore, even though we have shown that the combination of several metrics improved the prediction of PAF, little effort was placed on the optimization of a classification model. Thus, further research should focus on the construction of models for optimizing the prediction of AF. In addition, even though we developed an algorithm to predict PAF, it was not tested against any rhythms other than sinus rhythm.

Moreover, the P-wave delineation process was not completely automated, as we had to manually adjust 26.2% of the records, specially so in patients susceptible to PAF, probably due to the higher P-wave morphological variability that might have affected the delineator's performance.

Finally, further prospective studies with a larger sample size and considering wider time intervals before the onset of AF would be desirable in future research. In addition, the progressive inclusion of other rhythms and types of arrhythmias would be useful for the validation and improvement of predictive methodologies.

5.5 Final Remarks

This chapter has presented a study on the prediction of PAF. The aim of this study was to identify subjects at risk of developing PAF, and to predict the imminent PAF onset in patients known to be at risk. We have implemented and improved a selection of previously proposed methodologies and we carried out the analysis mostly in an unsupervised way, using the P-wave Quality Index tool. We found P-wave time and morphological variability to be the best predictors of PAF, and that the combination of several variability metrics improved the prediction performance.

The tools used to carry this study are being made available as a Matlab® toolbox of resources at **http://github.com/diogotecelao/AFPrediction**. The toolbox automatically performs: ECG signal pre-processing, delineation, feature extraction, metric calculation, and statistical analysis of significance.

Chapter 6

Prediction of Postoperative Atrial Fibrillation

This chapter presents a study on the prediction of postoperative atrial fibrillation (POAF). The aim of this study was to predict POAF in clinically relevant scenario. This included predicting POAF with up to 48 hours of antecedence, without supervision, and in a real world clinical dataset. This was performed by extracting a selection of features and metrics from the electrocardiogram (ECG), and by limiting the analysis to high-quality P-waves using the P-wave Quality index tool (presented in Chapter 4). The ability of the extracted metrics to predict POAF was studied at the following timestamps prior to the arrhythmia onset: 1h, 2h, 4h, 6h, 12h, 18h, 24h, 30h, 36h, 42h, and 48h. Finally, POAF prediction was performed by combining several indices.

6.1 Introduction

Postoperative atrial fibrillation (POAF) is the most frequent complication of cardiac surgery [9]. Similarly to other types of atrial fibrillation (AF), it affects patient well-being by increasing the risk of stroke, hemodynamic compromise, and mortality [13, 14]. In addition, it increases the hospital length of stay, and incurs additional treatment costs [9]. In the long term, patients with an episode of POAF have a twofold increase in cardiovascular mortality [12], and a substantial increase in the risk of future AF and ischemic stroke, compared to patients who remained in sinus rhythm after surgery [13–15].

Prediction of POAF has been mostly studied using preoperative risk stratification models, aiming to identify patients at risk of developing POAF after the surgery. Those studies have used risk factors, preoperative ECG features, or a combination of the two (Chapter 3). Even though those studies obtained promising results, they were not sufficiently reliable and robust to be applied clinically. To this regard, the best study achieved

an accuracy of 0.83.

Most recently, research has focused on beat-to-beat analyses of the ECG during continuous monitoring, thus enabling real-time clinical prediction of AF. Overall, these studies have identified variability in P-wave time- and morphology features to be a predictor of paroxysmal AF (PAF). However, these methods have only been tested during the two hours prior to PAF and have not been tested in a broader control group containing heart rhythms other than the sinus rhythm.

In this chapter we have investigated the ability of the previously proposed methods to predict POAF up to 48 hours before its onset using a real world clinical database. The present study is the first testing the ability of variability methods to predict POAF, and is the first studying the prediction of AF so long before from the arrhythmia onset. Prediction of POAF was performed up to 48 hours before its onset using a combination of several methods, and mostly without supervision.

6.2 Methods and materials

6.2.1 Study population

The MIMIC-III (Medical Information Mart for Intensive Care III) [142], a freely accessible critical care database, was used. It contains clinical notes and physiological measurements from intensive care unit (ICU) patients at Beth Israel Deaconess Medical Center between 2001 and 2012. The dataset is divided into two separate parts: one containing clinical information (MIMIC-III clinical database; version 1.4), and the other containing matched waveform signals (MIMIC-III waveform database matched subset; version 1.0) for a subset of ICU stays. The clinical database includes information such as demographics, vital sign measurements made at the bedside (approximately 1 data point per hour), laboratory test results, procedures, and medications. The waveform database almost always includes one or more electrocardiogram (ECG) signals, and often included continuous photoplethysmogram (PPG) signals and arterial blood pressure (ABP) waveforms.

6.2.1.1 Cohort selection

A POAF group containing 19 patients was selected using the following inclusion criteria: 1) AF was noted in the patients' clinical notes; 2) the AF episode happened after cardiac or thoracic surgery; 3) there were at least 48 hours of ECG recordings before the first AF episode; and 4) the ECG signal exhibited P-waves during the period before the onset of POAF. Having selected a group of eligible patients, we manually annotated the exact time at which the first AF episode started, and extracted 48 hours of lead II ECG recordings leading up to the onset of POAF. From those 48 hours, we extracted one-hour segments starting at the following timestamps prior to the arrhythmia onset: 1h, 2h, 4h, 6h, 12h, 18h, 24h, 30h, 36h, 42h, and 48h (Figure 6.1). This resulted in 209 recordings.

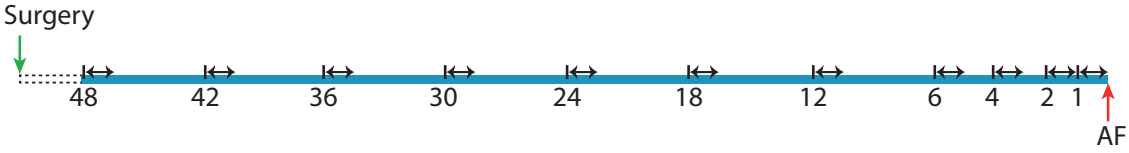

Figure 6.1: One-hour recordings were extracted at different timestamps relative to the onset of atrial fibrillation.

The control group contained 63 subjects and was selected using the following inclusion criteria: 1) had cardiac or thoracic surgery; 2) had no registered heart rhythm other than sinus rhythm, sinus tachycardia or sinus bradycardia; and 3) had at least 48 hours of ECG recordings with visible P-waves. The first 48 hours of available lead II ECG were then used, and one-hour segments were extracted at the same timestamps, but using the 48^{th} hour of ECG recording as a reference (instead of the onset of POAF). Finally, we excluded any one-hour recordings in which the median heart rate calculated over 1-minute segments was lower than 50 beats per minute (bpm) or higher than 110 bpm. This resulted in 530 recordings.

Finally, we used the following criteria to decide whether to include the one-hour recordings in the analysis: 1) at least 50% of the beats had P-waves of high-quality; 2) P-wave peaks could be well identified; and 3) P-wave onset and offset could be well delineated. While P-wave quality assessment was performed automatically, P-wave detection and delineation performance was assessed visually at random signal portions.

6.2.2 Signal preprocessing and delineation

Signal preprocessing and delineation of ECG fiducial points was performed in a similar manner to that described in Section 6.2, but with some differences. Firstly, baseline wander and high-frequency noise were removed with a band-pass filter. Next, R-peaks were detected using the Pan, Hamilton and Tompkins algorithm [100, 101], as in Section 5.2.2, with the exception that the algorithm was applied over 100s windows to reduce the effect of high amplitude artefacts. This was followed by a signal quality verification stage (see Section 6.2.2.1), which was added with the purpose of removing ECG beats which were highly corrupted by artefacts. Using the high-quality beats, we then performed Q- and P-wave delineation using the phasor transform, as described in Section 5.2.2. Finally, we excluded noise-corrupted P-waves using the P-wave quality index (PQI) tool (see Chapter 4), but applying the first decision stage over the entire one-hour recordings (i.e. the P-wave template from the first PQI decision stage was created using all the available P-waves). This further reduced the effect of any high amplitude noise that remained after exclusion of data using the signal quality verification algorithm.

6.2.2.1 ECG Signal Quality Index

Signal quality analysis was performed using a previously proposed template-matching algorithm [114]. The signal quality index (SQI) algorithm consisted of two parts. Firstly, the heart rate and the beat-to-beat intervals were derived from the R-peak detections. These were compared to a series of thresholds to determine whether their values were physiologically plausible. Secondly, a template beat was calculated as the average of each individual beat in a 10s window. The Pearson correlation between each individual beat and this template was calculated, and compared to an empirical threshold indicating the minimum acceptable correlation (0.66). A window was labelled as low quality if it failed any of these tests. The template-matching process is illustrated in Figure 6.2.

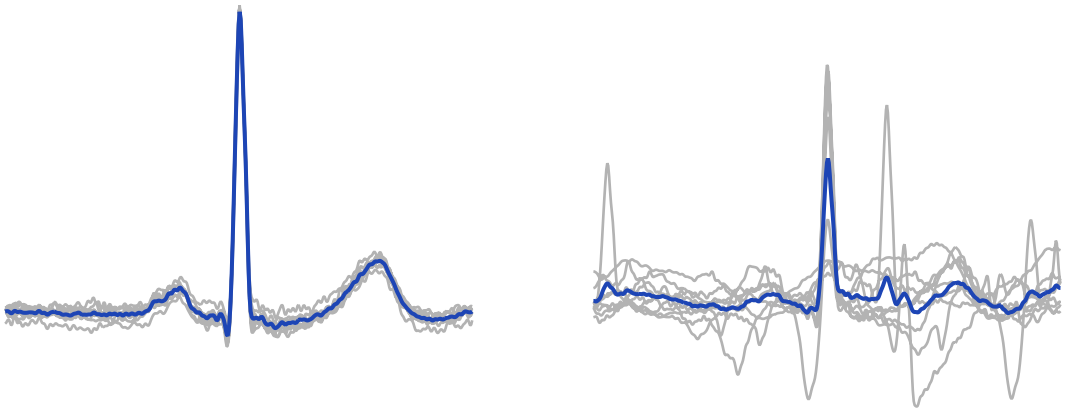

Figure 6.2: Template-matching for electrocardiogram (ECG) signal quality assessment. ECG signals were segmented into windows, and the correlation between individual beats (grey lines) in a window and the windows' beat template (blue lines) was calculated. If the correlation was bellow an empirically determined threshold then the segment was labelled as of low-quality (right). Otherwise, the segment was labelled as of high-quality (left). Adapted from Charlton *et al.* [119]

6.2.3 Feature extraction and metric calculation

The features presented in Section 5.2.3 were extracted from the ECG signal in a beat-to-beat fashion (Table 5.1). Those features were then summarised as metrics, using the same methods presented in Section 5.2.4. Briefly, several metrics were calculated, including the mean and standard deviation of the feature time-series, and the more complex calculation of linear and non-linear variability measurements. In addition, the heart rate variability (HRV) was also studied by extracting several common time and frequency indices from the RR series, as detailed in Section 5.2.4.

Linear variability was computed, firstly, by dividing a features' time-series into segments of S samples. Then, the variability within each segment was computed as the difference between the 10^{th} and 90^{th} percentiles. Finally, the variability time course was estimated using the slope (α) of a fitted linear model. Hence, a positive slope indicated

increasing variability, and vice-versa, while a null-slope indicated constant variability. In addition, variability was also quantified by taking the median value of the variability series. S was varied from 10 to 100.

Non-linear variability was quantified using the central tendency measurement (CTM). Briefly, the CTM was based on the plots of the d^{th}-order differences and was calculated as the proportion of points that fall within a circle of radius ρ around the origin. Hence, the higher the CTM, the lower the variability, and vice-versa. The lag/order was varied from 1 to 10 and the radius was varied from 0.1 to 10 multiples of the standard deviation of the analysed feature data-series.

6.2.4 Statistical analysis and performance assessment

We compared controls and POAF patients at each extracted timestamp (i.e. controls *vs.* POAF patients at a given hour before the arrhythmia onset) using non-parametric Wilcoxon rank sum tests (Mann-Whitney U-test). Results were considered as significant if their p-value was smaller than 0.01. A Holm-Sidak correction was performed at a significance level of 0.05.

The performance of each statistically significant metric (uncorrected $p < 0.01$) for discriminating between groups was assessed using the area under the receiver operating curve (AUC) statistic, evaluated through comparisons of the predicted labels with the records' true groups. Finally, prediction of POAF was performed using a decision tree model. All the metrics that had presented statistically significant differences, independently of the time-stamp in which they were significant, were considered as candidates to be included in the model. The stopping criterion used for the three growth was that each node contained fewer than 5% of all the observations. Performance was assessed using 10-fold cross validation, individually at each of the studied time-stamps. Given the imbalance between the two classes, only the sensitivity and specificity statistics were used, as they are independent of class distributions. During the analysis, POAF was considered the positive class.

6.3 Results

6.3.1 Inclusion of data

Out of the total 739 recordings from POAF patients and controls, 77 were excluded from the analysis. From the 77 recordings, 37.6% were (correctly) excluded due to excessive noise and artefacts, 10.3% due to the fact that the PQI tool excluded too many P-waves unnecessarily (P-wave template corrupted by noise), 42.8% because P-wave peak, onset or offset delineation was overall faulty (due to the nature of the signal), and 0.1% due to the absence of P-waves. The number of subjects which contributed data at each time-stamp varied and are presented in Figure 6.3. Finally, from the records that were included in the

Table 6.1: Clinical characteristics of the subjects that were included in the analysis.

	POAF patients	**Controls**	**Total**	**POAF *vs.* Controls**
Age	71.9 ± 9.2	58.8 ± 13.8	62.0 ± 14.0	Z=3.8, $p = 1.7 \times 10^{-4}$
Gender (Male/ Female)	11/ 8	34/ 25	45/ 33	χ^2=4.2×10^{-4}, $p = 0.9$
Surgery type (Cardiac/ Thoracic)	19/ 0	19/ 40	38/ 40	χ^2=26.4, $p = 2.7 \times 10^{-7}$

analysis, we had to manually adjust the P-wave peak search window in 1.4% of records, and the P-wave onset/offset search window in 3.3%.

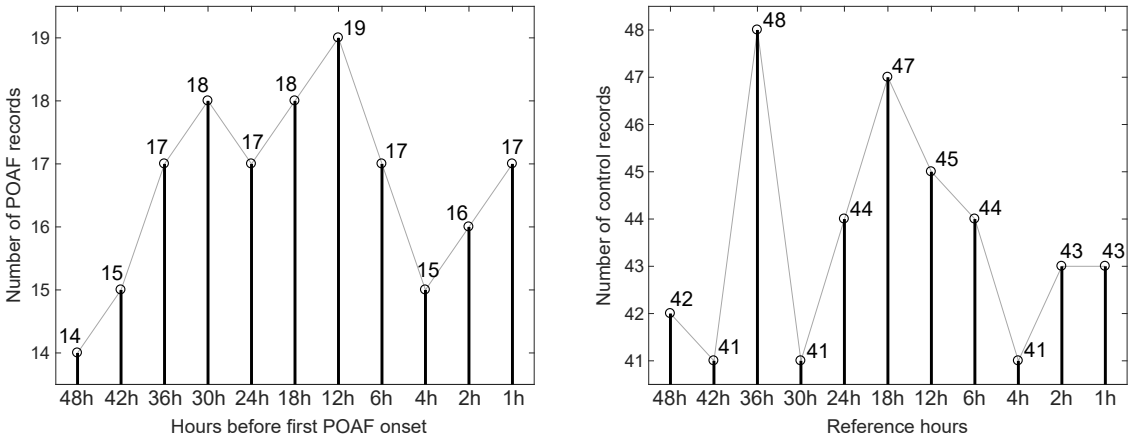

Figure 6.3: Number of records per time stamp used in the analysis for POAF patients (left) and controls (right).

Patients who had POAF were significantly ($p < 0.05$) older, and had a higher proportion of cardiac surgeries (Table 6.1).

6.3.2 Prediction of postoperative atrial fibrillation

6.3.2.1 Simple statistical metrics

Table 6.2 shows the significant results (uncorrected $p < 0.01$) obtained when comparing controls and POAF patients at the several timestamps using the simple statistical metrics. The mean and standard deviation of several time and morphological features was greater in POAF patients than in controls. Only the standard deviation of $PQ_{\text{on RRnorm}}$ was significant when correcting for multiple comparisons (corrected $p < 0.05$).

6.3.2.2 Linear variability

Significant results obtained when testing the linear variability metrics are presented in Table 6.3. Briefly, the variability in both time and morphological features were significantly different between POAF and controls far away from the arrhythmia (at least 18 hours

Table 6.2: Significant results (uncorrected $p < 0.01$) obtained when testing the simple statistical metrics. No results remained significant after correction for multiple comparisons. The median value (and the first/ third quartiles) of each group are presented.

Metrics	Controls	POAF	*P*-value	AUC
1 hour before POAF				
SD(PQ_{on}) [ms]	16.182 (14.550 / 18.468)	19.483 (16.121 / 23.838)	7.5×10^{-3}	0.72
SD($PQ_{on\ RRnorm}$) [a.u.]	0.025 (0.022 / 0.028)	0.032 (0.026 / 0.037)	4.8×10^{-3}	0.74
SD($PR_{peak\ RRnorm}$) [ms]	0.013 (0.011 / 0.018)	0.020 (0.014 / 0.027)	9.5×10^{-3}	0.72
2 hours before POAF				
SD($P_{dur.\ RRnorm}$) [a.u.]	0.024 (0.021 / 0.029)	0.032 (0.026 / 0.035)	4.8×10^{-3}	0.74
SD($PQ_{on\ RRnorm}$) [a.u.]	0.023 (0.020 / 0.027	0.032 (0.024 / 0.036)	1.6×10^{-3}	0.77
SD($PR_{on\ RRnorm}$) [a.u.]	0.020 (0.018 / 0.025)	0.029 (0.023 / 0.037)	1.9×10^{-3}	0.77
SD($PR_{peak\ RRnorm}$) [a.u.]	0.011 (0.009 / 0.017)	0.020 (0.014 / 0.025)	5.9×10^{-3}	0.74
6 hours before POAF				
SD(PQ_{on}) [ms]	16.392 (14.262 / 18.707)	18.756 (16.919 / 26.307)	5.8×10^{-3}	0.73
SD($PQ_{on\ RRnorm}$) [a.u.]	0.025 (0.020 / 0.029)	0.033 (0.027 / 0.038)	2.4×10^{-4}	0.81
SD($PQ_{off\ RRnorm}$) [a.u.]	0.016 (0.012 / 0.020)	0.021 (0.015 / 0.029)	9.8×10^{-3}	0.72
SD(PR_{peak}) [ms]	5.980 (4.436 / 7.571)	8.992 (6.061 / 13.023)	5.5×10^{-3}	0.73
SD(PR_{peak}) [ms]	0.011 (0.008 / 0.014)	0.017 (0.011 / 0.026)	4.5×10^{-3}	0.74
18 hours before POAF				
$\overline{\mathbf{P_{eucl.\ dist.}}}$ [a.u.]	0.108 (0.088 / 0.140)	0.145 (0.137 / 0.173)	9.3×10^{-3}	0.71
36 hours before POAF				
$\overline{\mathbf{P_{off\ amp.}}}$ [mV]	-0.084 (-0.108 / -0.068)	-0.056 (-0.083 / -0.046)	7.4×10^{-3}	0.72
$\overline{\mathbf{PQ_{level}}}$ [mV]	-0.079 (-0.099 / -0.063)	-0.056 (-0.071 / -0.044)	4.3×10^{-3}	0.74
$\overline{\mathbf{Q_{on\ amp.}}}$ [mV]	-0.072 (-0.094 / -0.057)	-0.048 (-0.063 / -0.037)	6.4×10^{-3}	0.72
SD($P_{amp.}$) [mV]	0.017 (0.011 / 0.029)	0.026 (0.020 / 0.046)	9.6×10^{-3}	0.71
42 hours before POAF				
SD($PR_{off\ RRnorm}$) [a.u.]	0.013 (0.010 / 0.017)	0.018 (0.015 / 0.020)	7.3×10^{-3}	0.74
SD($PQ_{off\ RRnorm}$) [a.u.]	0.015 (0.011 / 0.019)	0.019 (0.016 / 0.024)	8.2×10^{-3}	0.73
48 hours before POAF				
$\overline{\mathbf{P_{eucl.\ dist.}}}$ [a.u.]	0.098 (0.085 / 0.014)	0.147 (0.125 / 0.180)	8.3×10^{-3}	0.74

SD, standard deviation.

away), while only variability in P-wave time features were significant as the arrhythmia onset got closer. The POAF group had higher median variability for all the presented timestamps, and showed increasing variability (slope, α) in all timestamps except 18 hours before the arrhythmia onset. In contrast, controls presented decreasing variability in the majority of variability regression slope results. The PQ interval (duration and level) was the only feature which had metrics significant after correction for multiple comparisons at 6 and 48 hours before the arrhythmia onset.

6.3.2.3 Non-linear variability

Non-linear CTM was found to significantly differ between POAF patients and controls in all the tested timestamps (Table 6.4). Controls were associated with higher CTM in the great majority of features and timestamps. However, CTM applied over the features $P_{\text{vel. disp.}}$, $P_{\text{fin. dur}}$, $P_{\text{magn.}}$, PQ_{level}, $PQ_{\text{level Pnorm}}$ and $P_{\text{off amp.}}$ was sometimes higher in POAF patients. Finally, several P-wave time- and morphological metrics were significant after correction for multiple comparisons 1, 12, 18, 42 and 48 hours before the arrhythmia onset, all of them showing higher CTM in controls.

6.3.2.4 Heart rate variability

Several HRV metrics were significantly different between POAF patients and controls either close (up to 2 hours) or far (at least 42 hours) from the arrhythmia onset (Table 6.5). All the presented time domain metrics were higher in POAF, whilst all the frequency domain metrics, with the exception of pHF, were higher in controls. Only pLF, pHF, and the LF/HF ratio were significant after correction for multiple corrections during the last hour before AF onset.

6.3.2.5 Prediction performance

A decision tree was built (Figure 6.4) using the significant metrics listed above, and was tested at each time-stamp (Figure 6.5 and Table 6.6). Its sensitivity increased as the arrhythmia got closer, while its specificity was approximately constant. POAF was predicted 48 hours before the arrhythmia onset with a sensitivity of 0.74 and a specificity of 0.70. The decision tree used metrics derived from mean feature calculations and non-linear variability.

Table 6.3: Significant results (uncorrected $p < 0.01$) obtained when using the linear variability metrics. Results significant after correction for multiple corrections (corrected $p < 0.05$) are shaded in grey. The median value (and the first/ third quartiles) of each group are presented.

Metrics	Controls	POAF	*P*-value	AUC
1 hour before POAF ($S = 5$)				
$m(\mathbf{PQ_{on\ RRnorm}})$ [a.u.]	4.8×10^{-2} (4.0×10^{-2} / 5.7×10^{-2})	5.8×10^{-2} (5.0×10^{-2} / 7.6×10^{-2})	4.1×10^{-3}	0.74
$m(\mathbf{PR_{peak\ RRnorm}})$ [a.u.]	1.5×10^{-2} (1.2×10^{-2} / 2.2×10^{-2})	2.2×10^{-2} (2.0×10^{-2} / 4.0×10^{-2})	3.9×10^{-3}	0.74
2 hours before POAF ($S = 10$)				
$m(\mathbf{P_{dur.\ RRnorm}})$ [a.u.]	5.5×10^{-2} (4.6×10^{-2} / 6.5×10^{-2})	7.2×10^{-2} (5.8×10^{-2} / 8.1×10^{-2})	2.9×10^{-3}	0.75
$m(\mathbf{PQ_{on\ RRnorm}})$ [a.u.]	5.1×10^{-2} (4.3×10^{-2} / 6.1×10^{-2})	7.6×10^{-2} (5.7×10^{-2} / 8.4×10^{-2})	1.1×10^{-3}	0.78
$m(\mathbf{PR_{on\ RRnorm}})$ [a.u.]	4.0×10^{-2} (3.6×10^{-2} / 5.1×10^{-2})	6.5×10^{-2} (4.8×10^{-2} / 7.3×10^{-2})	4.1×10^{-3}	0.75
$m(\mathbf{PR_{peak\ RRnorm}})$ [a.u.]	1.7×10^{-2} (1.3×10^{-2} / 3.0×10^{-2})	3.5×10^{-2} (2.2×10^{-2} / 4.6×10^{-2})	3.9×10^{-3}	0.75
4 hours before POAF ($S = 10$)				
$m(\mathbf{PR_{peak\ RRnorm}})$ [a.u.]	1.7×10^{-2} (1.3×10^{-2} / 2.4×10^{-2})	2.8×10^{-2} (1.9×10^{-2} / 4.1×10^{-2})	8.2×10^{-3}	0.73
6 hours before POAF ($S = 10$)				
$m(\mathbf{P_{fin.\ dur.}})$ [ms]	24.0 (16.0 / 30.0)	28.0 (24.0 / 37.0)	5.7×10^{-3}	0.73
$m(\mathbf{P_{dur.\ RRnorm}})$ [a.u.]	5.9×10^{-2} (5.3×10^{-2} / 6.3×10^{-2})	7.1×10^{-2} (6.1×10^{-2} / 9.2×10^{-2})	2.2×10^{-3}	0.76
$m(\mathbf{PQ_{on}})$ [ms]	38.0 (32.0 / 44.0)	44.0 (36.0 / 58.0)	5.4×10^{-3}	0.73
$m(\mathbf{PQ_{on\ RRnorm}})$ [a.u.]	5.6×10^{-2} (4.4×10^{-2} / 6.2×10^{-2})	7.9×10^{-2} (6.0×10^{-2} / 9.0×10^{-2})	2.9×10^{-4}	0.80
$m(\mathbf{PR_{peak}})$ [ms]	8.0 (8.0 / 12.0)	16.0 (11.0 / 21.0)	1.6×10^{-3}	0.75
$m(\mathbf{PR_{on\ RRnorm}})$ [a.u.]	4.8×10^{-2} (3.7×10^{-2} / 5.4×10^{-2})	6.9×10^{-2} (4.8×10^{-2} / 8.7×10^{-2})	5.3×10^{-3}	0.73
$m(\mathbf{PR_{peak\ RRnorm}})$ [a.u.]	1.5×10^{-2} (1.3×10^{-2} / 2.5×10^{-2})	3.3×10^{-2} (1.9×10^{-2} / 5.4×10^{-2})	4.5×10^{-3}	0.74
12 hours before POAF ($S = 5$)				
$\alpha(\mathbf{Q_{on\ amp.}})$ [mV]	-2.0×10^{-6} (-6.1×10^{-6} / -3.0×10^{-6})	6.8×10^{-6} (-1.1×10^{-6} / 1.1×10^{-5})	5.2×10^{-3}	0.72
18 hours before POAF ($S = 10$)				
$\alpha(\mathbf{PQ_{on}})$ [ms]	9.9×10^{-4} (-8.9×10^{-3} / 6.9×10^{-3})	-9.5×10^{-3} (-2.4×10^{-2} / -2.0×10^{-3})	1.4×10^{-3}	0.76
$\alpha(\mathbf{PQ_{level,\ Pnorm}})$ [mV]	2.5×10^{-4} (-7.2×10^{-5} / 1.5×10^{-3})	-4.1×10^{-4} (-1.1×10^{-2} / 8.1×10^{-5})	1.3×10^{-3}	0.78
$\alpha(\mathbf{PR_{on}})$ [ms]	-1.7×10^{-4} (-8.5×10^{-3} / 7.0×10^{-3})	-7.1×10^{-3} (-2.8×10^{-2} / -1.7×10^{-3})	4.8×10^{-3}	0.73
$\alpha(\mathbf{PR_{peak}})$ [ms]	2.5×10^{-4} (-3.5×10^{-3} / 4.4×10^{-3})	-5.5×10^{-3} (-1.5×10^{-2} / -1.7×10^{-4})	8.1×10^{-3}	0.71
36 hours before POAF ($S = 30$)				
$\alpha(\mathbf{P_{min.\ vel.}})$ [mV]	-7.7×10^{-7} (-5.7×10^{-6} / 9.5×10^{-6})	1.9×10^{-5} (7.9×10^{-7} / 4.3×10^{-5})	7.7×10^{-3}	0.72
$\alpha(\mathbf{P_{vel.\ disp.}})$ [mV]	-1.1×10^{-6} (-1.1×10^{-5} / 1.6×10^{-5})	2.7×10^{-5} (2.4×10^{-6} / 6.2×10^{-5})	9.2×10^{-3}	0.71
$\alpha(\mathbf{P_{energy}})$ [μV^2]	-7.1×10^{-10} (-6.3×10^{-9} / 7.4×10^{-9})	1.1×10^{-8} (3.0×10^{-9} / 2.3×10^{-8})	3.4×10^{-3}	0.74
$\alpha(\mathbf{P_{energy\ norm.}})$ [V/s]	-1.1×10^{-10} (-6.9×10^{-10} / 4.7×10^{-10})	1.1×10^{-9} (5.1×10^{-10} / 1.8×10^{-9})	1.6×10^{-3}	0.76
$\alpha(\mathbf{P_{gauss.\ error}})$ [a.u.]	-5.1×10^{-8} (-3.8×10^{-6} / 2.4×10^{-6})	4.5×10^{-6} (-7.1×10^{-8} / 7.9×10^{-6})	4.9×10^{-3}	0.73
$m(\mathbf{P_{fin.\ dur.}})$ [ms]	24.0 (16.0 / 28.0)	32.0 (24.0 / 37.0)	6.7×10^{-3}	0.72
42 hours before POAF ($S = 5$)				
$m(\mathbf{P_{fin.\ dur.}})$ [ms]	16.0 (16.0 / 24.0)	24.0 (24.0 / 32.0)	8.8×10^{-3}	0.72
$m(\mathbf{PQ_{off}})$ [ms]	24.0 (16.0 / 24.0)	24.0 (24.0 / 32.0)	1.0×10^{-2}	0.72
$m(\mathbf{PQ_{on\ RRnorm}})$ [a.u.]	4.1×10^{-2} (3.6×10^{-2} / 5.4×10^{-2})	5.6×10^{-2} (4.4×10^{-2} / 6.9×10^{-2})	6.9×10^{-3}	0.74
$m(\mathbf{PQ_{off\ RRnorm}})$ [a.u.]	3.1×10^{-2} (2.0×10^{-2} / 4.1×10^{-2})	4.2×10^{-2} (3.3×10^{-2} / 5.2×10^{-2})	4.8×10^{-3}	0.75
$m(\mathbf{PR_{off\ RRnorm}})$ [a.u.]	2.4×10^{-2} (1.4×10^{-2} / 3.5×10^{-2})	3.5×10^{-2} (3.0×10^{-2} / 4.3×10^{-2})	7.7×10^{-3}	0.73
48 hours before POAF ($S = 10$)				
$\alpha(\mathbf{P_{magn.}})$ [mV]	-1.4×10^{-5} (-3.4×10^{-5} / -8.1×10^{-7})	6.7×10^{-6} (-5.2×10^{-6} / 1.6×10^{-5})	2.1×10^{-3}	0.78
$\alpha(\mathbf{P_{rms\ norm.}})$ [mV]	-6.3×10^{-7} (-1.5×10^{-6} / -4.1×10^{-9})	2.7×10^{-7} (7.8×10^{-8} / 6.4×10^{-7})	9.3×10^{-3}	0.73
$\alpha(\mathbf{P_{energy}})$ [μV^2]	-4.7×10^{-9} (-1.8×10^{-8} / 1.3×10^{-9})	2.0×10^{-9} (1.5×10^{-10} / 9.2×10^{-9})	7.0×10^{-3}	0.74
$\alpha(\mathbf{P_{energy\ norm.}})$ [V/s]	-4.1×10^{-10} (-1.2×10^{-9} / 4.1×10^{-11})	2.1×10^{-10} (9.6×10^{-12} / 1.6×10^{-9})	3.9×10^{-3}	0.76
$\alpha(\mathbf{PQ_{level,\ Pnorm}})$ [mV]	-3.7×10^{-3} (-7.0×10^{-2} / -6.6×10^{-4})	3.0×10^{-5} (-2.1×10^{-4} / 1.8×10^{-3})	7.4×10^{-4}	0.86
$m(\mathbf{P_{eucl.\ dist.}})$ [a.u.]	9.3×10^{-2} (7.3×10^{-2} / 1.2×10^{-1})	1.4×10^{-1} (9.6×10^{-2} / 1.9×10^{-1})	6.6×10^{-3}	0.74

m, mean variability; α, slope of regression.

Table 6.4: Significant results (uncorrected $p < 0.01$) obtained when testing the non-linear variability results. The presented values of optimal ρ correspond to the number of multiples of the standard deviation which minimised the p-value for the correspondent feature. Results significant after correction for multiple corrections (corrected $p < 0.05$) are shaded in grey. The median value (and the first/ third quartiles) of each group are presented (in arbitrary units).

Metrics	ρ	Controls	POAF	P-values	AUC
1 hour before POAF (lag= 5)					
$\mathbf{PQ_{off\ RRnorm}}$	3.0	0.8944 (0.8751 / 0.9085)	0.9089 (0.8996 / 0.9367)	2.7×10^{-3}	0.75
$\mathbf{PR_{off\ RRnorm}}$	8.0	1.0000 (0.9994 / 1.0000)	0.9994 (0.9979 / 1.0000)	6.0×10^{-3}	0.71
$\mathbf{PR_{on\ RRnorm}}$	4.0	0.9863 (0.9791 / 0.9931)	0.9741 (0.9554 / 0.9814)	3.1×10^{-4}	0.80
$\mathbf{PR_{peak\ RRnorm}}$	3.5	0.9849 (0.9677 / 0.9942)	0.9529 (0.9303 / 0.9735)	5.4×10^{-4}	0.79
$\mathbf{P_{area}}$	13.0	1.0000 (1.0000 / 1.0000)	1.0000 (0.9999 / 1.0000)	1.2×10^{-3}	0.62
$\mathbf{P_{dur.\ RRnorm}}$	4.0	0.9822 (0.9758 / 0.9884)	0.9714 (0.9637 / 0.9767)	3.3×10^{-4}	0.80
$\mathbf{P_{eucl.\ dist.}}$	4.0	0.9906 (0.9822 / 0.9930)	0.9754 (0.9658 / 0.9834)	3.1×10^{-5}	0.85
$\mathbf{P_{gauss.\ A}}$	6.0	0.9981 (0.9958 / 0.9990)	0.9947 (0.9907 / 0.9977)	3.9×10^{-3}	0.74
$\mathbf{P_{magn.}}$	4.0	0.9826 (0.9770 / 0.9896)	0.9721 (0.9564 / 0.9816)	9.2×10^{-4}	0.78
$\mathbf{P_{area\ norm.}}$	4.0	0.9776 (0.9687 / 0.9843)	0.9647 (0.9560 / 0.9724)	4.8×10^{-4}	0.79
$\mathbf{P_{rms\ norm.}}$	5.0	0.9912 (0.9855 / 0.9935)	0.9842 (0.9787 / 0.9896)	5.0×10^{-3}	0.73
$\mathbf{P_{vel.\ disp.}}$	4.0	0.9821 (0.9781 / 0.9870)	0.9764 (0.9726 / 0.9816)	3.5×10^{-3}	0.74
CCI	10.5	1.0000 (1.0000 / 1.0000)	1.0000 (0.9993 / 1.0000)	3.5×10^{-3}	0.74
2 hours before POAF (lag= 8)					
$\mathbf{PQ_{off}}$	8.5	1.0000 (1.0000 / 1.0000)	1.0000 (0.9987 / 1.0000)	2.8×10^{-3}	0.67
$\mathbf{PR_{on}}$	11.0	1.0000 (1.0000 / 1.0000)	1.0000 (0.9992 / 1.0000)	5.2×10^{-3}	0.65
$\mathbf{PR_{peak}}$	14.0	1.0000 (1.0000 / 1.0000)	1.0000 (0.9991 / 1.0000)	5.9×10^{-3}	0.67
$\mathbf{PQ_{off\ RRnorm}}$	9.5	1.0000 (1.0000 / 1.0000)	1.0000 (0.9991 / 1.0000)	6.7×10^{-3}	0.65
$\mathbf{P_{eucl.\ dist.}}$	4.5	0.9952 (0.9892 / 0.9968)	0.9884 (0.9756 / 0.9942)	2.8×10^{-3}	0.76
$\mathbf{P_{gauss.\ C}}$	14.5	1.0000 (0.9992 / 1.0000)	0.9991 (0.9986 / 0.9997)	7.1×10^{-3}	0.71
$\mathbf{P_{area\ norm.}}$	3.5	0.9593 (0.9466 / 0.9684)	0.9372 (0.9327 / 0.9548)	5.3×10^{-3}	0.74
$\mathbf{P_{fin.\ dur.}}$	6.0	0.9994 (0.9981 / 1.0000)	1.0000 (0.9998 / 1.0000)	2.7×10^{-3}	0.75
4 hours before POAF (lag= 6)					
$\mathbf{P_{dur\ RRnorm}}$	6.0	0.9995 (0.9988 / 1.0000)	0.9982 (0.9957 / 0.9995)	4.6×10^{-3}	0.75
$\mathbf{PR_{on\ RRnorm}}$	4.5	0.9928 (0.9882 / 0.9966)	0.9856 (0.9667 / 0.9913)	4.4×10^{-3}	0.75
$\mathbf{P_{area}}$	13.0	1.0000 (1.0000 / 1.0000)	1.0000 (0.9995 / 1.0000)	1.5×10^{-3}	0.65
6 hours before POAF (lag= 6)					
$\mathbf{P_{energy}}$	11.5	0.9991 (0.9981 / 0.9997)	0.9980 (0.9961 / 0.9987)	8.3×10^{-3}	0.72
$\mathbf{P_{eucl.\ dist.}}$	4.5	0.9941 (0.9895 / 0.9968)	0.9911 (0.9816 / 0.9922)	4.8×10^{-3}	0.74
$\mathbf{P_{fin.\ dur.}}$	6.0	0.9993 (0.9976 / 1.0000)	1.0000 (0.9994 / 1.0000)	5.9×10^{-3}	0.72
12 hours before POAF (lag= 4)					
$\mathbf{PQ_{off}}$	13.0	1.0000 (1.0000 / 1.0000)	1.0000 (0.9990 / 1.0000)	5.2×10^{-4}	0.65
$\mathbf{PQ_{off\ RRnorm}}$	9.5	1.0000 (1.0000 / 1.0000)	1.0000 (0.9991 / 1.0000)	2.2×10^{-4}	0.64
$\mathbf{PQ_{on}}$	7.5	1.0000 (1.0000 / 1.0000)	0.9998 (0.9990 / 1.0000)	7.6×10^{-4}	0.71
$\mathbf{PQ_{on\ RRnorm}}$	8.0	1.0000 (1.0000 / 1.0000)	0.9997 (0.9985 / 1.0000)	6.1×10^{-4}	0.72

Continued on next page

Table 6.4 – continued from previous page

Metrics	ρ	Controls	POAF	*P*-values	AUC
$\mathbf{PQ_{level}}$	11.0	0.9996 (0.9990 / 1.0000)	1.0000 (1.0000 / 1.0000)	6.4×10^{-3}	0.70
$\mathbf{PR_{off}}$	16.0	1.0000 (1.0000 / 1.0000)	1.0000 (0.9994 / 1.0000)	2.5×10^{-3}	0.62
$\mathbf{PR_{off\ RRnorm}}$	11.0	1.0000 (1.0000 / 1.0000)	1.0000 (0.9993 / 1.0000)	7.0×10^{-3}	0.63
$\mathbf{PR_{on}}$	8.0	1.0000 (1.0000 / 1.0000)	0.9995 (0.9986 / 1.0000)	5.6×10^{-5}	0.75
$\mathbf{PR_{on\ RRnorm}}$	7.5	1.0000 (0.9991 / 1.0000)	0.9981 (0.9971 / 0.9998)	1.4×10^{-3}	0.74
$\mathbf{PR_{peak}}$	13.0	1.0000 (1.0000 / 1.0000)	1.0000 (0.9987 / 1.0000)	2.4×10^{-3}	0.66
$\mathbf{PR_{peak\ RRnorm}}$	22.0	1.0000 (1.0000 / 1.0000)	1.0000 (1.0000 / 1.0000)	7.2×10^{-3}	0.58
$\mathbf{P_{al}}$	7.0	1.0000 (1.0000 / 1.0000)	0.9998 (0.9997 / 1.0000)	6.2×10^{-4}	0.70
$\mathbf{P_{dur.}}$	7.0	1.0000 (1.0000 / 1.0000)	0.9998 (0.9997 / 1.0000)	6.2×10^{-4}	0.70
$\mathbf{P_{eucl.\ dist.}}$	5.0	0.9974 (0.9937 / 0.9990)	0.9938 (0.9902 / 0.9953)	3.4×10^{-4}	0.79
$\mathbf{P_{magn.}}$	10.0	0.9996 (0.9989 / 1.0000)	1.0000 (1.0000 / 1.0000)	3.2×10^{-3}	0.72
$\mathbf{P_{vel.\ disp.}}$	10.5	1.0000 (0.9993 / 1.0000)	1.0000 (1.0000 / 1.0000)	8.1×10^{-3}	0.67
18 hours before POAF (lag= 9)					
$\mathbf{PQ_{off}}$	9.5	1.0000 (1.0000 / 1.0000)	1.0000 (0.9988 / 1.0000)	6.5×10^{-3}	0.63
$\mathbf{P_{al}}$	1.0	0.2750 (0.2431 / 0.3175)	0.2420 (0.2335 / 0.2543)	2.6×10^{-3}	0.74
$\mathbf{P_{dur.}}$	1.0	0.2750 (0.2431 / 0.3175)	0.2421 (0.2335 / 0.2543)	2.6×10^{-3}	0.74
$\mathbf{P_{dur.\ RRnorm}}$	1.0	0.2907 (0.2503 / 0.3414)	0.2464 (0.2320 / 0.2735)	9.3×10^{-3}	0.71
$\mathbf{P_{eucl.\ dist.}}$	4.0	0.9901 (0.9862 / 0.9931)	0.9827 (0.9633 / 0.9870)	1.2×10^{-3}	0.76
$\mathbf{P_{ini.\ dur.}}$	2.0	0.6733 (0.6358 / 0.7047)	0.6468 (0.6180 / 0.6552)	8.5×10^{-3}	0.71
24 hours before POAF (lag= 6)					
$\mathbf{P_{al}}$	8.0	1.0000 (1.0000 / 1.0000)	1.0000 (1.0000 / 1.0000)	4.9×10^{-3}	0.59
$\mathbf{P_{dur.}}$	8.0	1.0000 (1.0000 / 1.0000)	1.0000 (1.0000 / 1.0000)	4.9×10^{-3}	0.59
$\mathbf{P_{magn.}}$	3.0	0.9461 (0.9349 / 0.9627)	0.9270 (0.9185 / 0.9405)	7.8×10^{-3}	0.72
$\mathbf{P_{max.\ vel.}}$	3.5	0.9629 (0.9585 / 0.9706)	0.9576 (0.9467 / 0.9605)	3.9×10^{-3}	0.74
$\mathbf{P_{min.\ vel.}}$	3.5	0.9625 (0.9561 / 0.9697)	0.9548 (0.9508 / 0.9633)	7.1×10^{-3}	0.72
30 hours before POAF (lag= 7)					
$\mathbf{P_{area}}$	12.5	1.0000 (1.0000 / 1.0000)	1.0000 (1.0000 / 1.0000)	8.3×10^{-3}	0.58
$\mathbf{WI}_t$	9.0	1.0000 (1.0000 / 1.0000)	1.0000 (1.0000 / 1.0000)	2.1×10^{-3}	0.61
36 hours before POAF (lag= 3)					
$\mathbf{PQ_{level}}$	6.5	0.9972 (0.9952 / 0.9993)	0.9936 (0.9907 / 0.9966)	2.1×10^{-3}	0.75
$\mathbf{PQ_{level\ Rnorm}}$	6.5	0.9972 (0.9951 / 0.9989)	0.9950 (0.9911 / 0.9974)	9.5×10^{-3}	0.71
$\mathbf{P_{area}}$	5.5	0.9980 (0.9957 / 0.9993)	0.9945 (0.9908 / 0.9983)	9.2×10^{-3}	0.71
$\mathbf{P_{area\ norm.}}$	4.5	0.9886 (0.9833 / 0.9938)	0.9806 (0.9763 / 0.9886)	7.7×10^{-3}	0.72
42 hours before POAF (lag= 1)					
$\mathbf{PQ_{on}}$	3.5	0.9387 (0.9278 / 0.9454)	0.9459 (0.9429 / 0.9520)	9.6×10^{-3}	0.73
$\mathbf{P_{energy\ norm.}}$	12.5	0.9993 (0.9986 / 0.9998)	0.9982 (0.9973 / 0.9987)	1.1×10^{-3}	0.79
$\mathbf{P_{fin.\ dur.}}$	3.5	0.9402 (0.9251 / 0.9537)	0.9571 (0.9470 / 0.9654)	4.6×10^{-3}	0.75
$\mathbf{WI}_t$	8.0	1.0000 (1.0000 / 1.0000)	0.9998 (0.9997 / 1.0000)	9.7×10^{-3}	0.68
48 hours before POAF (lag= 7)					
$\mathbf{PQ_{level\ Pnorm}}$	17.5	0.9982 (0.9975 / 0.9989)	0.9993 (0.9989 / 1.0000)	3.9×10^{-3}	0.76
$\mathbf{P_{area}}$	3.0	0.9181 (0.9010 / 0.9338)	0.8926 (0.8859 / 0.9262)	9.3×10^{-3}	0.73
$\mathbf{P_{gauss.\ W}}$	20.0	1.0000 (1.0000 / 1.0000)	1.0000 (0.9989 / 1.0000)	9.9×10^{-3}	0.66

Continued on next page

Table 6.4 – continued from previous page

Metrics	ρ	Controls	POAF	*P*-values	AUC
$\mathbf{P_{magn.}}$	2.5	0.9210 (0.8766 / 0.9576)	0.8810 (0.8358 / 0.9132)	7.4×10^{-3}	0.74
$\mathbf{P_{area\ norm.}}$	2.0	0.8111 (0.7547 / 0.8485)	0.7089 (0.6585 / 0.7424)	1.5×10^{-4}	0.84
$\mathbf{P_{rms\ norm.}}$	2.0	0.7889 (0.7472 / 0.8466)	0.6877 (0.6503 / 0.7236)	1.2×10^{-3}	0.79
$\mathbf{P_{off\ amp.}}$	8.0	0.9971 (0.9952 / 0.9986)	0.9992 (0.9980 / 1.0000)	2.3×10^{-3}	0.77
$\mathbf{P_{fin.\ dur.}}$	5.5	0.9988 (0.9962 / 0.9995)	0.9997 (0.9993 / 1.0000)	9.3×10^{-3}	0.73
$\mathbf{WI}_t$	2.5	0.8301 (0.8163 / 0.8425)	0.8083 (0.8008 / 0.8246)	3.7×10^{-3}	0.76

Table 6.5: Significant results (uncorrected $p < 0.01$) obtained when performing the heart rate variability analysis. Results significant after correction for multiple corrections (corrected $p < 0.05$) are shaded in grey. The median value (and the first/ third quartiles) of each group are presented.

Metrics	Controls	POAF	*P*-value	AUC
1 hour before POAF				
SDSD [ms]	0.016 (0.011 / 0.026)	0.036 (0.021 / 0.052)	7.9×10^{-3}	0.72
RMSSD [ms]	0.016 (0.011 / 0.026)	0.036 (0.021 / 0.052)	7.9×10^{-3}	0.72
pLF [%]	52.156 (48.477 / 55.595)	46.672 (39.238 / 49.735)	1.8×10^{-3}	0.76
pHF [%]	47.844 (44.405 / 51.523)	53.328 (50.265 / 60.762)	1.8×10^{-3}	0.76
LF/HF [a.u.]	1.090 (0.941 / 1.252)	0.875 (0.646 / 0.989)	1.8×10^{-3}	0.76
2 hours before POAF				
SDSD [ms]	0.014 (0.010 / 0.026)	0.036 (0.018 / 0.058)	3.6×10^{-3}	0.75
RMSSD [ms]	0.014 (0.010 / 0.026)	0.036 (0.018 / 0.058)	3.6×10^{-3}	0.75
HF [ms^2]	0.469 (0.382 / 0.611)	1.050 (0.528 / 1.540)	4.5×10^{-3}	0.74
pLF [%]	53.194 (47.798 / 55.715)	46.132 (41.862 / 49.871)	5.6×10^{-3}	0.74
pHF [%]	46.806 (44.285 / 52.202)	53.868 (50.129 / 58.138)	5.6×10^{-3}	0.74
LF/HF [a.u.]	1.136 (0.916 / 1.258)	0.856 (0.721 / 0.995)	5.6×10^{-3}	0.74
42 hours before POAF				
NN50 [a.u.]	7.000 (1.000 / 31.000)	29.000 (11.500 / 115.250)	7.4×10^{-3}	0.74
pNN50 [a.u.]	0.001 (0.000 / 0.006)	0.006 (0.002 / 0.023)	9.5×10^{-3}	0.73
48 hours before POAF				
VLF [ms^2]	3.894 (3.483 / 4.427)	3.445 (3.152 / 3.638)	4.7×10^{-3}	0.76

Table 6.6: Classification results obtained when predicting postoperative at the several tested timestamps.

	Hours before POAF										
	48h	**42h**	**36h**	**30h**	**24h**	**18h**	**12h**	**6h**	**4h**	**2h**	**1h**
Sensitivity	0.74	0.77	0.74	0.74	0.81	0.82	0.92	0.83	0.87	0.88	0.89
Specificity	0.70	0.79	0.79	0.74	0.72	0.79	0.76	0.76	0.75	0.79	0.78

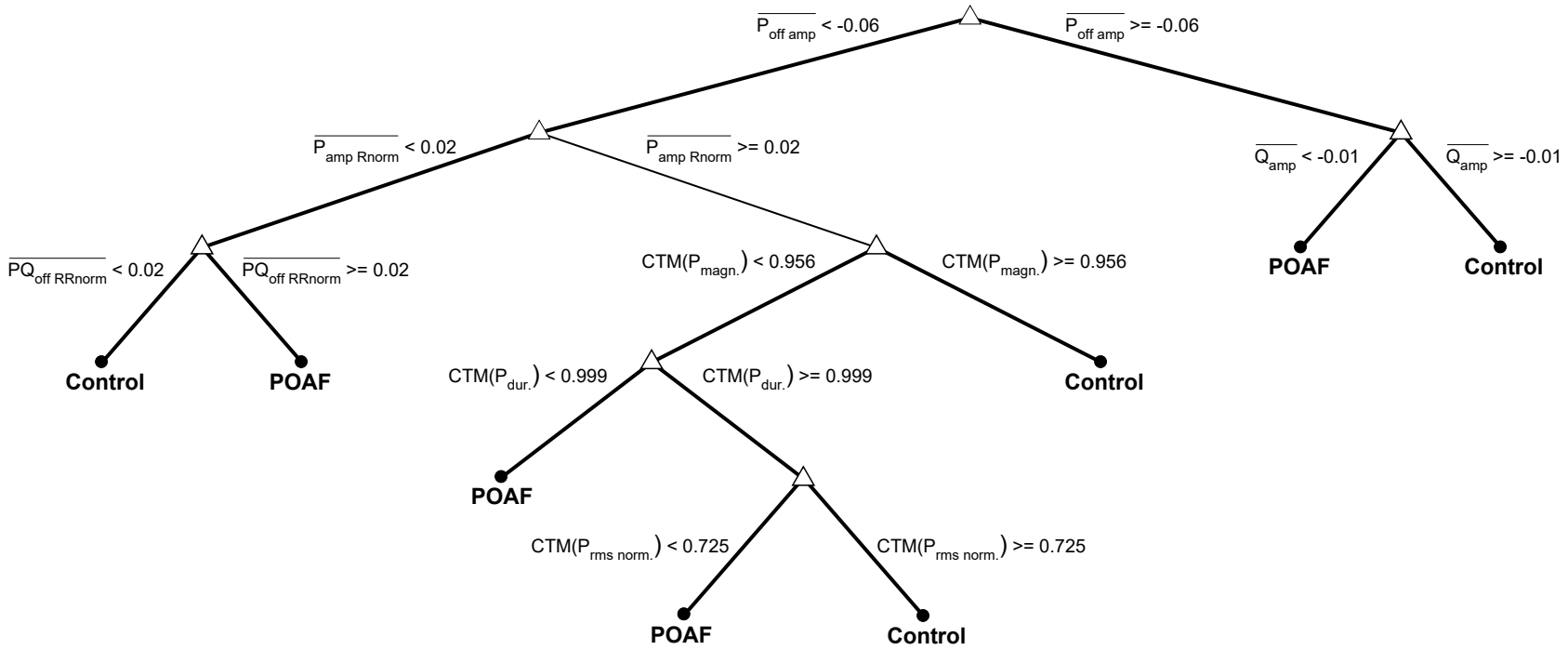

Figure 6.4: Decision tree used to predict postoperative atrial fibrillation. This model used metrics from mean feature calculations and non-linear variability (central tendency measurement; CTM) variability.

6.4 Discussion

POAF is the most common complication of cardiac surgery. It can be life-threatening, increases hospital length of stay, and entails additional treatment costs. Prophylactic treatment has been shown to be effective, and may be beneficial if the patients at risk of POAF could be identified early enough.

In the present study we aimed to predict the occurrence of POAF up to 48 hours before its onset by analysing ECG signals from a real world clinical dataset, without supervision. Thus, this study has aimed to predict POAF in a scenario similar to that found in the clinical practice.

We found several P-wave and HRV metrics to be predictive of POAF, even far away from the arrhythmia onset. P-wave variability was again the best predictor, as it provided the highest number of significant features both close and far from the arrhythmia. Overall, we found patients who developed POAF to have greater and increasing P-wave variability, while controls had less and decreasing variability (Figure 6.6). The analysis was conducted without supervision for the great majority of records, except for 4.7%, where the P-wave delineation processes needed to be manually adjusted. POAF was predicted by combining metrics derived from mean feature calculations and non-linear variability.

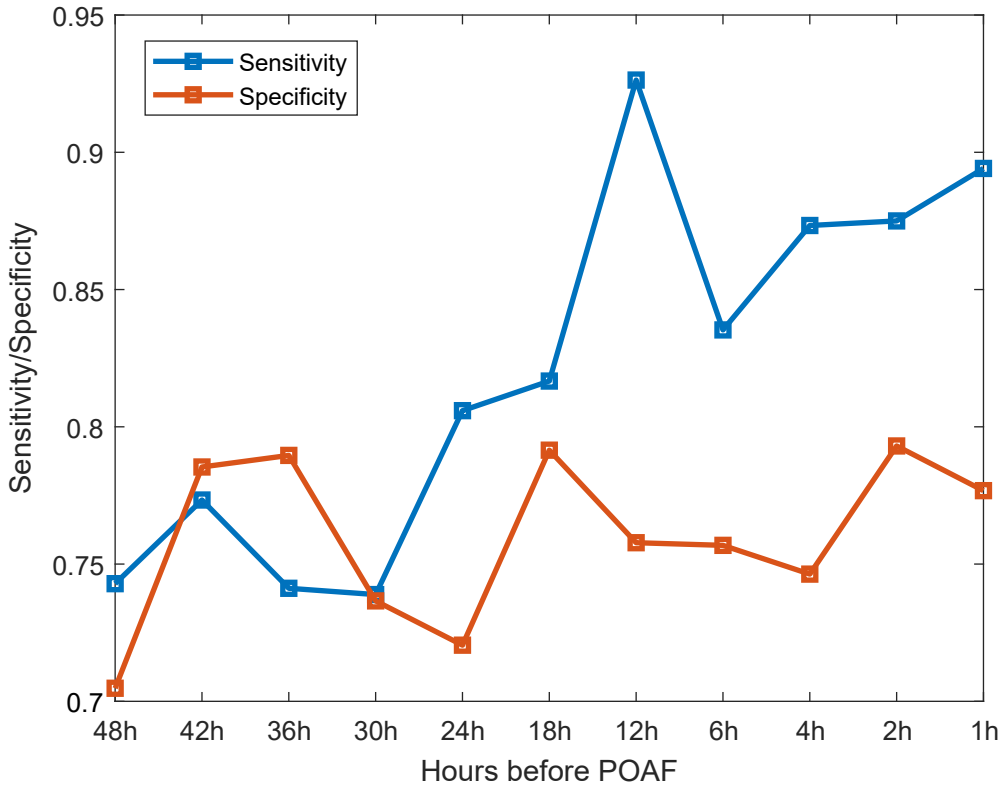

Figure 6.5: Sensitivity and specificity obtained when trying to predict postoperative atrial fibrillation. The ability to predict POAF increased as the arrhythmia got closer.

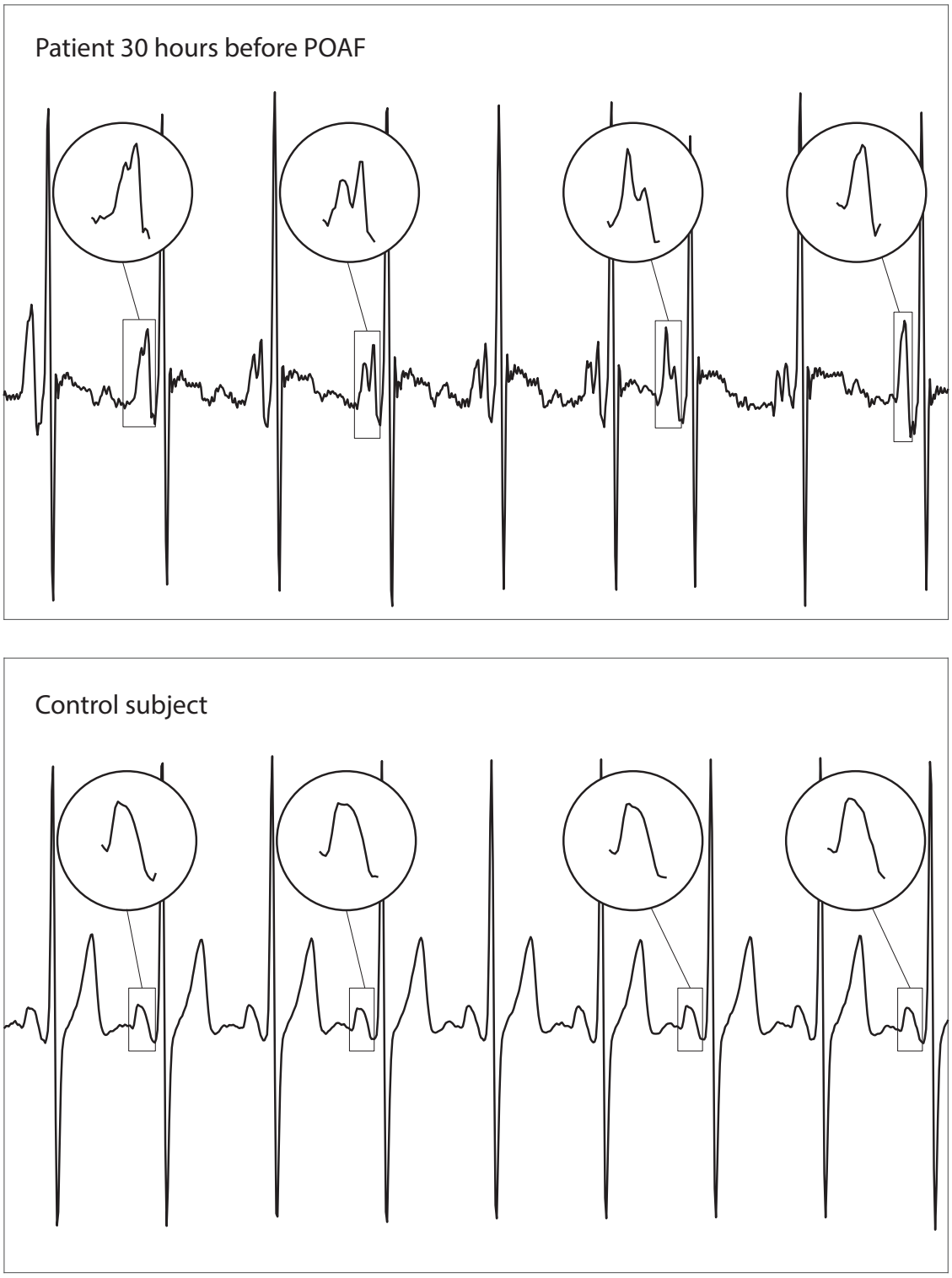

Figure 6.6: Example of representative electrocardiogram signals corresponding to a patient 30 hours before the onset of postoperative atrial fibrillation (POAF; top) and a control (bottom). Patients who develop POAF are characterized with higher variability in P-wave morphology, as depicted.

6.4.1 Prediction of postoperative atrial fibrillation

The ability to predict POAF increased as the arrhythmia got closer (Figure 6.5), reflecting the increasing prominence of atrial alterations that predispose to and trigger the arrhythmia. Our model was able to predict POAF 48 hours before its onset with good sensitivity (0.74) and specificity (0.70). Even though our results were not validated in an independent dataset, performance was assessed using 10-fold cross validation. This allowed to minimise overfitting and to get insight on how the model performed in an independent dataset.

Our prediction results are comparable to those from other studies (with accuracies ranging from 0.70 to 1.00), even though slightly worst (Table 3.3). However, it must be noted that the present study has aimed to predict POAF in a much more complex scenario: 1) we have considered a much longer time before onset of AF (48 hours), while most studies have only studied times close to the arrhythmia (2 hours, mostly); and 2) our method has rejected unreliable P-waves without supervision, while most studies excluded noisy P-waves by hand.

It is interesting to note that the decision tree used to predict POAF (Figure 6.4) used metrics from mean calculations and non-linear variability, while the decision trees used to predict PAF used *only* variability metrics (Figures 5.10 and 5.11). However, it should be remarked that the prediction of PAF and POAF should probably be performed separately (i.e., using different models), because they arise from two different contexts. For example, surgical patients are submitted to several intra- and postoperative factors (Figure 2.4) that dispose to POAF, which are not present in daily life. In addition, cardiac surgery causes a reduction in the activity of the autonomic nervous system, that is also not present in daily life [143].

6.4.2 P-wave variability in the postoperative setting

P-wave time- and morphological variability was found to be the best predictor of POAF, by providing the greatest number of significant metrics across all the tested timestamps. This is in line with our previous study (Chapter 5), where linear and non-linear variability were also found to be the best predictors of PAF.

Generally, patients who had POAF exhibited higher and increasing variability in several time and morphological features of the P-wave, throughout the studied 48 hours. It is also interesting that controls showed decreasing variability for most of the significant variability regression metrics. The changes which occur during recovery from surgery may explain these results. All surgical patients are subjected to several intra- and postoperative factors which dispose to POAF (Figure 2.4). However, controls and POAF patients seem to respond differently to such stimulae: while controls recover from it, showing decreasing variability, patients who develop POAF seem not to, presenting the higher and increasing variability which eventually leads to the arrhythmia.

Finally, it should be remarked that, unexpectedly, some metrics indicated decreased or decreasing variability in POAF patients, while the control group showed the reverse effect. Further studies are needed to understand if these results were false positives, or reflect a physiological change. Nonetheless, all the results which were significant when correcting for multiple corrections showed the expected effect of increased and increasing variability in POAF patients.

6.4.3 Heart rate variability

The HRV analysis revealed mostly a high number of autonomic changes close to the arrhythmia start. Indeed, the 2 hours preceding AF have been identified as a critical time where short-term autonomic changes influencing the trigger of this arrhythmia are likely to occur [139]. There were also relevant alterations in HRV markers a fairly long time before the arrhythmia (at least 42 hours). All these metrics showed good predictive value, as measured using the AUC.

Our results are in line with previous studies showing POAF patients to have: 1) increased SDRR, RMSSD and pNN50 in either thoracic [139] or coronary artery bypass grafting (CABG) [144] surgery; and 2) increased HF and decreased LF/HF ratio in both surgery types [88, 139, 145]. In opposition to our findings, there are also associations of increased LF [139] and LF/HF ratio [144] in patients who developed POAF. However, it is important to note that these differences may arise due to other factors that may differ between studies, such as the surgery type, disease states, medications, circadian rhythms and respiratory rate [139]. It is also interesting to note how these results of increased heart rate complexity in patients who developed POAF are in opposition to the findings in PAF, where complexity is often decreased, which may suggest different triggering mechanisms. Finally, even though POAF patients were significantly older than controls, our results of increased HRV in that group are not likely explained by age, given that HRV decreases with ageing [146].

6.4.4 Data inclusion

One important characteristic of a predicting tool aimed for use in a real world clinical setting is its ability to work without supervision. In the present work we aimed to predict POAF in an unsupervised fashion. There are at least two factors that affect the accuracy of AF prediction: noise, which leads to a high number of false alarms, and faulty delineations, which cause erroneous feature measurements.

Noise was addressed by combining an ECG signal quality index and the P-wave quality index tool, which worked well for the great majority of recordings. However, from the records that were excluded from the analysis, 10.3% were unnecessarily removed because the P-wave template created during the P-wave quality analysis was corrupted by noise, which affected the quality assessment, and led to discarding those records. Indeed,

the impact of excessive noise on the P-wave template creation has been identified as a limitation of that tool (Subsection 4.4.1).

Furthermore, 4.7% of the records included in the analysis needed their P-wave delineations to be manually optimised. Otherwise, P-wave features and metrics would have been wrongly estimated, affecting the prediction of AF. Future improvements in P-wave delineation would help to mitigate this issue. For instance, comparison of the P-wave signal (i.e. between the delineated onset and offset) with a Gaussian function or with a P-wave template could help to identify abnormal delineations. Moreover, identification of implausible P-wave feature variations could also be helpful. Currently, one P-wave delineator compares each delineated P-wave with a reference P-wave, and excludes them if they present a large difference (>25%) in a set of extracted features [147]. Applying such decision rules to detect faulty delineations, could result in modifications in the delineation parameters, improving the delineation.

6.4.5 Limitations and future work

Even though the proposed predictive model has achieved good classification performance, it was still not reliable enough to be implemented and used. Classification ability may be further improved with the combination of POAF risk factors and preoperative ECG features, which have shown good predictive value on their own (Table 3.2). Furthermore, the use of more complex classification models might improve the ability to predict AF.

In addition, as discussed, several other aspects still need to be considered so that the prediction of POAF might be performed in the clinical practice:

1. Improvement in the signal quality assessment, so that noise does not affect the subsequent stages of the algorithm, such as P-wave template creation;

2. Optimization of P-wave delineation for all possible variations in P-wave morphology. An optimal algorithm would detect failures in delineation and automatically adjust its parameters.

3. Study of a broader control group, including all heart rhythms other than AF and atrial flutter.

Finally, this study presents some limitations that merit consideration. Firstly, the sample of POAF patients was small (19 subjects). Larger prospective studies are therefore needed to further validate our findings and our predictive model. Secondly, the two groups were not matched with regard to patient age and surgery type. Finally, the ECG recordings from the two groups were also not matched with regard to the time since surgery. Nonetheless, our study represents a more real scenario where controls started being monitored at a random time after the surgery.

6.5 Final Remarks

This chapter has presented a study on the prediction of POAF. The aim of this study was to predict POAF in a clinically relevant scenario: with up to 48 hours of antecedence and without supervision. Thus, this study aimed to predict POAF in a scenario similar to that found in the clinical practice. Prediction of POAF was studied using the methods presented in Chapter 5, and at the following timestamps prior to the arrhythmia onset: 1h, 2h, 4h, 6h, 12h, 18h, 24h, 30h, 36h, 42h, and 48h. Several P-wave and HRV metrics were predictive of POAF, even far away from the arrhythmia onset. P-wave variability was again the best predictor, regardless of the time until the arrhythmia onset. Prediction of POAF was performed, mostly without supervision, up to 48 hours before its onset with with good sensitivity (0.74) and specificity (0.70).

Chapter 7

Conclusion

This chapter concludes the thesis by presenting an overview of its achievements, and recommended directions for future research.

7.1 Summary of thesis achievements

The aim of this thesis was to provide a proof of concept that postoperative atrial fibrillation (POAF) can be predicted using the electrocardiogram (ECG), with sufficient warning time for prophylactic treatments to be used. Chapters 3 to 6 presented investigations to address these aims. The achievements of the thesis are summarized as follows:

Chapter 3: State of the art on the prediction of postoperative atrial fibrillation

Chapter 3 presented a review of techniques previously proposed to predict POAF. These methods were divided into preoperative methods, which aim to identify patients at risk of developing POAF, and postoperative methods, designed to predict the imminent onset of atrial fibrillation (AF) using continuous monitoring.

The review indicated that preoperative risk stratification models are of clinical interest, but of limited applicability, given that they are not sufficiently reliable (the highest accuracy was 0.83). Moreover, novel studies were identified that perform beat-to-beat analyses of the ECG during continuous monitoring with promising results (accuracy ranged from 0.70 to 1.00). These studies found that variability in the P-wave (the wave that reflects atrial depolarization) time- and morphological features is predictive of paroxysmal AF.

The findings of the literature review were used to identify key areas for future research, which were addressed during this thesis. The identified directions for future

research were: 1) to develop techniques that ensure that the measurements extracted from signals are reliable; 2) to perform a comprehensive comparison of predictive methods, and to test whether combining methods improves the prediction of AF; and 3) to assess the performance of these algorithms in a scenario that is clinically relevant for the prediction of POAF.

Chapter 4: Assessment of P-wave quality

Chapter 4 presented and assessed the performance and utility of a novel tool that assesses P-wave quality. The aim of developing this tool was to ensure that only high quality and reliable P-waves were used in the AF prediction analyses.

The tool consisted of a two-stage algorithm which used P-wave template matching to assess quality. Its performance and utility was assessed using the AFPDB, a database of wearable ECG signals acquired from both healthy subjects and patients susceptible to paroxysmal AF. The algorithm's quality assessments of 97,989 P-waves were compared to manual annotations. The algorithm identified high quality P-waves with high sensitivity (0.93) and good specificity (0.82). The tool's clinical utility was assessed using: 1) mean absolute error comparisons of several well-known P-wave measurements (such as duration and amplitude) with and without the tool; and 2) evaluation of the performance of AF prediction by using metrics extracted from the P-waves with and without the tool. The tool improved the precision of feature measurements, and improved the prediction of AF, especially when using methods that analyse P-wave variability.

This tool will be especially useful with the future implementation of techniques that predict AF during long term monitoring using wearable sensors. This tool was used in the AF prediction analyses in Chapters 5 and 6.

Chapter 5: Prediction of paroxysmal atrial fibrillation

Chapter 5 presented a study on the prediction of paroxysmal AF, following the approach of the 2001 Computers in Cardiology challenge. The aim of this challenge was to identify subjects at risk of developing AF, and to predict the imminent onset of AF in patients known to be at risk. In this study, a selection of previously proposed methods were compared, and we investigated whether combining methods improved prediction of AF. This study used ECG recordings from 31 healthy subjects, 14 AF patients far away from the arrhythmia onset, and 15 AF patients just before arrhythmia onset.

Prediction of AF was performed as follows. Firstly, P-waves were delineated, and their quality was assessed automatically using the tool presented in Chapter 4. Secondly, a set of features was extracted from each high quality P-wave. Then, each time-series of feature measurements was summarised as a single metric using a selection of methods. Extracted metrics ranged from simple statistical indices such as the mean and standard deviation, to more complex measurements of variability, including heart rate variability

(HRV). Finally, the predictive ability of those metrics was tested individually and in combination.

Variability in P-wave time- and morphological features was found to be the best predictor of AF. Indeed, similarly to previous studies, we found that P-wave variability is greater in patients at risk of AF, and increases as the arrhythmia gets closer. Moreover, the autonomic nervous system was found to be altered in patients at risk of AF, as indicated by the HRV analysis. Finally, the combination of several metrics improved the prediction of AF. We could identify patients at risk of AF and to predict the imminent onset with good sensitivity (0.82) and high specificity (0.93). Performance might be further improved with the use of more complex classification models.

The low sampling frequency used in this study (128Hz) might have hindered accurate P-wave characterization, and consequently might have attenuated subtle and predictive P-wave modifications. Given that ECG devices used in the clinical setting commonly have low sampling frequencies, this study was useful for identifying metrics capable of predicting AF in such a context.

Further prospective studies with a larger sample size, considering wider time intervals before the onset of AF, including other heart rhythms and types of arrhythmias, and using commonly found sampling frequencies will be useful for the validation of these predictive techniques.

Chapter 6: Prediction of postoperative atrial fibrillation

Chapter 6 presented a study on the prediction of POAF using the MIMIC-III database. This study investigated the ability of the previously proposed methods to predict POAF up to 48 hours before its start, in a real world clinical scenario. This study was the first testing the ability of the variability methods to predict POAF. Furthermore, this study was the first studying the prediction of AF so long before from the arrhythmia onset (48 hours). Prediction of POAF was performed as in Chapter 5. This study included records from 48 controls and 19 patients who developed POAF.

P-wave variability was again the best predictor of AF, regardless of the time until the arrhythmia onset. Patients who developed POAF had greater and increasing P-wave variability, whilst controls had less and decreasing variability. This might indicate that while controls recover (decreasing variability) from the several intra- and postoperative factors that lead to POAF, the condition of patients who develop POAF seems to deteriorate further (increasing variability), leading to the arrhythmia. The autonomic nervous system was also found to have a role in POAF, both close (< 2 hours) and far (> 42 hours) from the arrhythmia, as indicated by the HRV analysis.

The combination of several metrics allowed POAF to be predicted 48 hours before its onset with good sensitivity (0.74) and specificity (0.70). Performance improved as the arrhythmia got closer. Even though these results were slightly worse than those from previous studies, this study aimed to predict POAF in a more complex scenario: 1) 48

hours away from the arrhythmia, whereas previous studies have mostly investigated the 2 hours before AF; and 2) mostly without supervision. Finally, automated P-wave quality assessment and delineation was good for the great majority of records, but not for a small minority.

Further studies with a larger sample size, and with the progressive inclusion of more heart rhythms are warranted. Further improvements in automated signal quality assessment and P-wave delineation will also be important in order to maximise the performance of AF prediction techniques, especially when used in the clinical setting.

7.2 Future work

Even though this work has addressed several aspects that are important for the clinical integration of AF prediction techniques, the performance of these methods is still suboptimal. Several aspects that need further improvement are now discussed.

7.2.1 Reliability of P-wave measurements

The prediction of AF is highly reliant on the extraction of measurements from P-waves. However, P-waves are of low-amplitude, making them difficult to delineate, and are highly susceptible to noise. Measurements derived from a wrongly delineated or noise-corrupted P-wave are often erroneous, reducing the accuracy of the prediction methods. Hence, making sure that only high quality and well delineated P-waves are used in the analysis is of utmost importance. Even though we have addressed this issue during this thesis, there are still ways in which this can be improved further.

P-wave delineation

During this thesis, we tested two different P-wave delineators (see Appendix A), and chose to use one based on the phasor transform because it performed better in ECG recordings from subjects susceptible to AF. However, during our analyses (Chapters 5 and 6) the parameters of the delineator had to be adjusted so that the delineation of some records was acceptable and reliable. Even though this was performed only to a small minority of records, this manual step is not feasible in the clinical setting, especially because the number of patients in hospital and severities of illnesses are increasing, while staffing levels are decreasing [119].

Hence, further work should focus on the improvement of P-wave delineation. One way in which this can be done is with the addition of verification stages. For instance, comparison of the P-wave signal (i.e, between the delineated onset and offset) with a Gaussian function or with a P-wave template could help to identify substantially abnormal delineations (e.g., P-wave onset in the middle of a fragmented P-wave, or P-wave offset in the Q-wave). Furthermore, identification of implausible P-wave feature variations could

also be helpful. To this end, one recent P-wave delineator compares each delineated P-wave with a reference P-wave, and excludes them if they present a large difference (>25%) in a set of extracted features [147]. In addition, the signal quality index (SQI) algorithm used in Chapter 6 also compares the RR intervals with a series of thresholds to determine whether their values are physiologically plausible [114]. Applying such decision rules to detect faulty P-wave delineations, could result in changes in the delineation parameters (e.g., in the P-wave peak search window), improving delineations.

P-wave quality assessment

Even though the P-wave quality assessment (PQI) tool presented in this thesis identified low quality P-waves with good performance, its performance on the MIMIC-III database was not perfect. Indeed, during that analysis, out of the 77 hour-long records that were removed due to excessive noise, 10.3% were unnecessarily excluded because P-wave templates were affected by noise, affecting the quality assessment. Even though this happened to a low percentage of records, clinically, this would delay or prevent the prediction of AF (at least by one hour, the record's duration).

Hence, further studies should improve the PQI tool so that it works well independently of the type of artefacts present in the records. A way in which this can be addressed is with the addition of a P-wave template verification stage. For instance, a template-verification stage where the obtained P-wave template could be compared with a Gaussian function, would safeguard the existence of a noise-corrupted P-wave template.

Furthermore, during the creation of template P-waves, it is assumed that the ECG recording has a single P-wave morphology, thus ignoring the possibility of a secondary P-wave morphology. Indeed, a recent study has found an increased number of P-waves matching a distinct secondary morphology to be significantly associated with paroxysmal AF, possibly indicating different conduction routes on the atrial myocardium [120]. If those secondary morphologies significantly differ from the main morphology of the created template, they will probably be classified as low quality and will consequently be discarded from the analysis. Thus, future improvements in P-wave quality assessment could include the possibility of an additional P-wave morphology with, for example, the creation of two different P-wave templates that could then be used to extract P-wave features and assess its quality.

7.2.2 Prediction of atrial fibrillation

Even though the methods used to predict AF obtained good and promising results, their performance was still insufficient for implementation in the clinical setting. There exist at least four ways in which prediction of POAF can be improved in the future:

1. Inclusion of risk factors (such as age, gender, medications) and preoperative ECG

features to improve performance. The combination of preoperative risk stratification and postoperative prediction models has not been reported. Given that these techniques achieve good performance separately, their combination could be of interest. In addition, variability analyses of preoperative ECG signals have never been performed and might also be of interest.

2. Development of other methods to measure the variability of the extracted P-wave features. In this regard, methods used to study HRV might be of interest.

3. Development of more complex classification models that can improve the performance to predict AF (e.g., support vector machine classifiers). However, care should be taken in order to avoid overfitting.

4. Inclusion of other physiological monitoring techniques, such as the photoplethysmogram (PPG). The PPG is an optical measurement technique that detects blood volume changes in the microvascular bed of tissue, providing information on the heart and the vascular and autonomic nervous systems. The PPG is an useful cardiac arrhythmia detector as it is very sensitive to any irregularity of the pulse. This is particularly apparent in phenomena such as atrial or ventricular premature beats. Recent studies have shown the PPG signal to be a reliable tool to detect and diagnose AF rhythm [148–150].

In addition, future studies should keep addressing the prediction of POAF under scenarios that are clinically relevant. Future research should then:

1. Study the prediction of AF at a time where it is still possible to prevent the arrhythmia.

2. Include diverse heart rhythms. Even though atrial fibrillation is the most common heart rhythm disorder in the postoperative period, other heart rhythms and arrhythmias can occur, such as tachyarrhythmias and bradyarrhythmias [110]. It is important to include these rhythms in future studies to ensure that the prediction of POAF in the clinical setting is reliable. To date, studies have only included ECG recordings during sinus rhythm.

3. Study the prediction of AF without supervision and using signals obtained from continuous monitoring of patients.

4. Study the prediction of POAF following other surgery types. To date, most studies have only aimed to predict POAF after cardiac surgery.

Finally, even though several studies have addressed the prediction of AF using continuous ECG monitoring, they have analysed small populations (the largest study has included 46 patients with paroxysmal AF and 53 healthy subjects; Table 3.3). Hence, in order to develop reliable prediction techniques, they must be tested and validated in a

substantially larger population. Freely available databases containing physiological measurements from intensive unit care (ICU) patients, such as the MIMIC-III [142] database, will help mitigate this issue, as they usually contain data for large populations.

7.3 Application in clinical practice

Hospital patients recovering from major cardiac surgery are at high risk of POAF, which can be life-threatening. By predicting the POAF early enough, the development of the arrhythmia could be potentially prevented using prophylactic treatments, thus reducing risks and hospital costs. The prophylactic effect depends on the pharmacokinetics of the drugs used (e.g., administered dose, rate of absorption, and rate of elimination). The pharmacokinetic parameter we are most interested is the time of onset of antiarrhythmic effect, which varies from drug to drug. In this application, it would be optimal if the prophylactic treatments peaked their effect as fast as possible, ideally the furthest from POAF onset.

There are two prophylactic drugs recommended for POAF prevention: beta-blockers and amiodarone (Section 2.3.4). The anti-arrhythmic effect of beta-blockers peaks immediately when administered intravenously (IV), and after 1 to 1.5 hours when administered orally [57]. In contrast, the anti-arrhythmic effect of amiodarone peaks between 1 minute and 1.5 hours after IV administration [62]. When administered orally, the drug concentration peaks after 3 to 7 hours [63, 64], but the anti-arrhythmic effects only start after 2 to 3 days [62, 65].

The hypothesis of this thesis is that ECG signals can be analysed to predict the onset of POAF with sufficient warning time to allow prophylactic treatment to be administered. In this thesis, a tool was developed to predict POAF, showing good sensitivity (0.74) and specificity (0.70) 48 hours before its onset (Chapter 6). Hence, by applying this tool, one would have the next 47 hours to prevent the onset of AF using prophylactic treatments. Despite the existence of several therapeutic plans (e.g., administration of amiodarone 400mg firstly IV and four days later orally), the aforementioned numbers suggest that the time between predicting POAF and its onset is sufficiently large for the prophylactic treatments to take anti-arrhythmic effects and consequently enable the prevention of POAF.

Nonetheless, further studies are required in order to determine which prophylactic drug and therapeutic plan is the best suited for this particular scenario, where the drugs need to be necessarily administered postoperatively, and during the time between predicting POAF and its onset. In addition, the efficiency of those treatments might also be monitored using the ECG. For instance, if it was possible to detect that a given prophylactic treatment was not efficiently preventing POAF, the drug and/or the therapeutic plan could be altered and tailored.

Bibliography

[1] G. Y. Lip, C. M. Brechin, and D. A. Lane. "The global burden of atrial fibrillation and stroke: A systematic review of the epidemiology of atrial fibrillation in regions outside North America and Europe." In: *Chest* 142.6 (2012), pp. 1489–1498. DOI: 10.1378/chest.11-2888.

[2] A. Ahlsson, E. Fengsrud, L. Bodin, and A. Englund. "Postoperative atrial fibrillation in patients undergoing aortocoronary bypass surgery carries an eightfold risk of future atrial fibrillation and a doubled cardiovascular mortality." In: *European Journal of Cardio-thoracic Surgery* 37.6 (2010), pp. 1353–1359. DOI: 10.1016/j.ejcts.2009.12.033.

[3] S. F. Aranki, D. P. Shaw, D. H. Adams, R. J. Rizzo, G. S. Couper, M. VanderVliet, J. J. Collins, L. H. Cohn, and H. R. Burstin. "Predictors of Atrial Fibrillation After Coronary Artery Surgery: Current Trends and Impact on Hospital Resources." In: *Circulation* 94 (1996), pp. 390–397. DOI: https://doi.org/10.1161/01.CIR.94.3.390.

[4] L. L. Creswell, R. B. Schuessler, M. Rosenbloom, and J. L. Cox. "Hazards of postoperative atrial arrhythmias." In: *The Annals of Thoracic Surgery* 56.3 (1993), pp. 539–549. DOI: 10.1016/0003-4975(93)90894-N.

[5] W. H. Maisel, J. D. Rawn, and W. G. Stevenson. "Atrial Fibrillation after Cardiac Surgery." In: *Annals of Internal Medicine* 135.12 (2001), pp. 1061–1073. DOI: 10.7326/0003-4819-135-12-200112180-00010.

[6] J. P. Mathew, M. L. Fontes, I. C. Tudor, J. Ramsay, P. Duke, C. D. Mazer, P. G. Barash, P. H. Hsu, and D. T. Mangano. "A multicenter risk index for atrial fibrillation after cardiac surgery." In: *JAMA: the Journal of the American Medical Association* 291.14 (2004), pp. 1720–1729. DOI: 10.1001/jama.291.14.1720.

[7] R. P. Villareal, R. Hariharan, B. C. Liu, B. Kar, V. V. Lee, M. Elayda, J. A. Lopez, A. Rasekh, J. M. Wilson, and A. Massumi. "Postoperative atrial fibrillation and mortality after coronary artery bypass surgery." In: *Journal of the American College of Cardiology* 43.5 (2004), pp. 742–748. DOI: 10.1016/j.jacc.2003.11.023.

[8] C. W. Hogue, L. L. Creswell, D. D. Gutterman, and L. A. Fleisher. "Epidemiology, mechanisms, and risks: American College of Chest Physicians guidelines for the prevention and management of postoperative atrial fibrillation after cardiac surgery." In: *Chest* 128.2 Suppl (2005), 9S–16S. DOI: 10.1378/chest.128.2.

[9] N. Echahidi, P. Pibarot, G. O'Hara, and P. Mathieu. "Mechanisms, Prevention, and Treatment of Atrial Fibrillation After Cardiac Surgery." In: *Journal of the American College of Cardiology* 51.8 (2008), pp. 793–801. DOI: 10.1016/j.jacc.2007.10.043.

[10] D. Kaireviciute, A. Aidietis, and G. Y. Lip. "Atrial fibrillation following cardiac surgery: clinical features and preventative strategies." In: *European Heart Journal* 30.4 (2008), pp. 410–425. DOI: 10.1093/eurheartj/ehn609.

[11] G. Y. H. Lip, L. Fauchier, S. B. Freedman, I. Van Gelder, A. Natale, C. Gianni, S. Nattel, T. Potpara, M. Rienstra, H.-F. Tse, and D. A. Lane. "Atrial fibrillation." In: *Nature Reviews Disease Primers* 2 (2016), pp. 1–26. DOI: 10.1038/nrdp.2016.16.

[12] A. Ahlsson, L. Bodin, E. Fengsrud, and A. Englund. "Patients with postoperative atrial fibrillation have a doubled cardiovascular mortality." In: *Scandinavian Cardiovascular Journal* 43.5 (2009), pp. 330–336. DOI: 10.1080/14017430802702291.

[13] G. Mariscalco and K. G. Engström. "Atrial fibrillation after cardiac surgery: Risk factors and their temporal relationship in prophylactic drug strategy decision." In: *International Journal of Cardiology* 129.3 (2008), pp. 354–362. DOI: 10.1016/j.ijcard.2007.07.123.

[14] K. Phan, H. S. Ha, S. Phan, C. Medi, S. P. Thomas, and T. D. Yan. "New-onset atrial fibrillation following coronary bypass surgery predicts long-term mortality: A systematic review and meta-analysis." In: *European Journal of Cardio-thoracic Surgery* 48.6 (2015), pp. 817–824. DOI: 10.1093/ejcts/ezu551.

[15] A. Saxena, D. T. Dinh, J. A. Smith, G. C. Shardey, C. M. Reid, and A. E. Newcomb. "Usefulness of postoperative atrial fibrillation as an independent predictor for worse early and late outcomes after isolated coronary artery bypass grafting (multicenter australian study of 19,497 patients)." In: *American Journal of Cardiology* 109.2 (2012), pp. 219–225. DOI: 10.1016/j.amjcard.2011.08.033.

[16] J. D. Aasbo, A. T. Lawrence, K. Krishnan, M. H. Kim, and R. G. Trohman. "Amiodarone Prophylaxis Reduces Major Cardiovascular Morbidity and Length of Stay after Cardiac Surgery : A Meta-Analysis." In: *Annals of Internal Medicine* 143.5 (2005), pp. 327–336. DOI: 10.7326/0003-4819-143-5-200509060-00008.

[17] E. Crystal, S. J. Connolly, K. Sleik, T. J. Ginger, and S. Yusuf. "Interventions on prevention of postoperative atrial fibrillation in patients undergoing heart surgery: A meta-analysis." In: *Circulation* 106.1 (2002), pp. 75–80. DOI: 10.1161/01.CIR.0000021113.44111.3E.

[18] E. M. Davis, K. A. Packard, and D. E. Hilleman. "Pharmacologic prophylaxis of postoperative atrial fibrillation in patients undergoing cardiac surgery: beyond beta-blockers." In: *Pharmacotherapy* 30.7 (2010), 274e–318e. DOI: 10.1592/phco.30.7.749.

[19] K. A. Eagle, R. A. Guyton, R. Davidoff, F. H. Edwards, G. A. Ewy, T. J. Gardner, J. C. Hart, H. C. Herrmann, L. D. Hillis, A. M. Hutter, B. W. Lytle, R. A. Marlow, W. C. Nugent, T. A. Orszulak, E. M. Antman, S. C. Smith, J. S. Alpert, J. L. Anderson, D. P. Faxon, V. Fuster, R. J. Gibbons, G. Gregoratos, J. L. Halperin, L. F. Hiratzka, S. A. Hunt, A. K. Jacobs, and J. P. Ornato. "ACC/AHA 2004 Guideline Update for Coronary Artery Bypass Graft Surgery: Summary Article." In: *Journal of the American College of Cardiology* 44.5 (2004), pp. 1146–1154. DOI: 10.1016/j.jacc.2004.07.021.

[20] K. Jongnarangsin and H. Oral. "Postoperative Atrial Fibrillation." In: *Cardiology Clinics* 27.1 (2009), pp. 69–78. DOI: 10.1016/j.ccl.2008.09.011.

[21] J. Gu, J. J. Andreasen, J. Melgaard, S. Lundbye-Christensen, J. Hansen, E. B. Schmidt, K. Thorsteinsson, and C. Graff. "Preoperative Electrocardiogram Score for Predicting New-Onset Postoperative Atrial Fibrillation in Patients Undergoing Cardiac Surgery." In: *Journal of Cardiothoracic and Vascular Anesthesia* 31.1 (2017), pp. 69–76. DOI: 10.1053/j.jvca.2016.05.036.

[22] Y. Reckman and E. Creemers. "Circulating circles predict postoperative atrial fibrillation." In: *Journal of the American Heart Association* 7.2 (2018), pp. 1–4. DOI: 10.1161/JAHA.117.008261.

[23] S. Poli, V. Barbaro, P. Bartolini, G. Calcagnini, and F. Censi. "Prediction of atrial fibrillation from surface ECG: Review of methods and algorithms." In: *Annali dell'Istituto Superiore di Sanita* 39.2 (2003), pp. 195–203.

[24] S. Sovilj, A. Van Oosterom, G. Rajsman, and R. Magjarevic. "ECG-based prediction of atrial fibrillation development following coronary artery bypass grafting." In: *Physiological Measurement* 31.5 (2010), pp. 663–677. DOI: 10.1088/0967-3334/31/5/005.

[25] A. Martínez, R. Alcaraz, and J. J. Rieta. "Study on the P-wave feature time course as early predictors of paroxysmal atrial fibrillation." In: *Physiological Measurement* 33.12 (2012), pp. 1959–1974. DOI: 10.1088/0967-3334/33/12/1959.

[26] A. Martínez, R. Alcaraz, and J. J. Rieta. "Morphological variability of the P-wave for premature envision of paroxysmal atrial fibrillation events." In: *Physiological Measurement* 35.1 (2014), pp. 1–14. DOI: 10.1088/0967-3334/35/1/1.

[27] F. Censi, I. Corazza, E. Reggiani, G. Calcagnini, E. Mattei, M. Triventi, and G. Boriani. "P-wave Variability and Atrial Fibrillation." In: *Scientific Reports* 6.26799 (2016), pp. 1–7. DOI: 10.1038/srep26799.

[28] L. S. Lilly. *Pathophysiology of Heart Disease : a collaborative project of medical students and faculty*. 5th ed. Wolters Kluwer/Lippincott Williams & Wilkins, 2011, p. 461. ISBN: 1605477230.

[29] R. O. Bono, D. Mann, D. Zipes, and P. Libby. *Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine*. 9th ed. Elsevier Saunders, 2012, p. 2048. ISBN: 9781437727708.

[30] H. Stefansdottir, T. Aspelund, V. Gudnason, and D. O. Arnar. "Trends in the incidence and prevalence of atrial fibrillation in Iceland and future projections." In: *Europace* 13.8 (2011), pp. 1110–1117. DOI: 10.1093/europace/eur132.

[31] E. Anter, M. Jessup, and D. J. Callans. "Atrial fibrillation and heart failure: Treatment considerations for a dual epidemic." In: *Circulation* 119.18 (2009), pp. 2516–2525. DOI: 10.1161/CIRCULATIONAHA.108.821306.

[32] S. Colilla, A. Crow, W. Petkun, D. E. Singer, T. Simon, and X. Liu. "Estimates of current and future incidence and prevalence of atrial fibrillation in the U.S. adult population." In: *American Journal of Cardiology* 112.8 (2013), pp. 1142–1147. DOI: 10.1016/j.amjcard.2013.05.063.

[33] B. P. Krijthe, A. Kunst, E. J. Benjamin, G. Y. Lip, O. H. Franco, A. Hofman, J. C. Witteman, B. H. Stricker, and J. Heeringa. "Projections on the number of individuals with atrial fibrillation in the European Union, from 2000 to 2060." In: *European Heart Journal* 34.35 (2013), pp. 2746–2751. DOI: 10.1093/eurheartj/eht280.

[34] D. G. Sherman, L Goldman, R. B. Whiting, K Jurgensen, M Kaste, and J. D. Easton. "Thromboembolism in patients with atrial fibrillation." In: *Archives of Neurology* 41.7 (1984), pp. 708–710. DOI: 10.1001/archneur.1984.04050180030011.

[35] P. A. Wolf, R. D. Abbott, and W. B. Kannel. "Original Contributions Atrial Fibrillation as an Independent Risk Factor for Stroke : The Framingham Study." In: *Stroke* 22.8 (1991), pp. 983–988. DOI: 10.1161/01.STR.22.8.983.

[36] W. B. Kannel, R. D. Abbott, D. D. Savage, and P. M. McNamara. "Epidemiologic Features of Chronic Atrial Fibrillation: The Framingham Study." In: *New England Journal of Medicine* 306.5 (1982), pp. 1018–1022. DOI: 10.1056/NEJM198204293061703.

[37] H.-J. Lin, P. A. Wolf, M. Kelly-Hayes, A. S. Beiser, C. S. Kase, E. J. Benjamin, and R. B. D'Agostino. "Stroke Severity in Atrial Fibrillation." In: *Stroke* 27.10 (1996), pp. 1760–1764.

[38] L. Frost, H. Mølgaard, E. H. Christiansen, C. J. Jacobsen, H. Allermand, and P. E. Thomsen. "Low vagal tone and supraventricular ectopic activity predict atrial fibrillation and flutter after coronary artery bypass grafting." In: *European Heart Journal* 16.6 (1995), pp. 825–831. DOI: 10.1093/oxfordjournals.eurheartj.a061002.

[39] L. Frost, E. H. Christiansen, H. Mølgaard, C.-J. Jacobsen, H. Allermand, and P. E. B. Thomsen. "Premature atrial beat eliciting atrial fibrillation after coronary artery bypass grafting." In: *Journal of Electrocardiology* 28.4 (1995), pp. 297–305. DOI: 10.1016/S0022-0736(05)80047-0.

[40] T. Killip and J. H. Gault. "Mode of onset of atrial fibrillation in man." In: *American Heart Journal* 20.2 (1965), pp. 172–179. DOI: 10.1016/0002-8703(65)90064-5.

[41] C. Kolb, S. Nürnberger, G. Ndrepepa, B. Zrenner, A. Schömig, and C. Schmitt. "Modes of initiation of paroxysmal atrial fibrillation from analysis of spontaneously occurring episodes using a 12-lead Holter monitoring system." In: *American Journal of Cardiology* 88.8 (2001), pp. 853–857. DOI: 10.1016/S0002-9149(01)01891-4.

[42] T. J. Wang, M. G. Larson, D. Levy, R. S. Vasan, E. P. Leip, P. A. Wolf, R. B. D'Agostino, J. M. Murabito, W. B. Kannel, and E. J. Benjamin. "Temporal relations of atrial fibrillation and congestive heart failure and their joint influence on mortality: The Framingham heart study." In: *Circulation* 107.23 (2003), pp. 2920–2925. DOI: 10.1161/01.CIR.0000072767.89944.6E.

[43] A. Alonso, B. P. Krijthe, T. Aspelund, K. A. Stepas, M. J. Pencina, C. B. Moser, M. F. Sinner, N. Sotoodehnia, J. D. Fontes, A. C. J. Janssens, R. A. Kronmal, J. W. Magnani, J. C. Witteman, A. M. Chamberlain, S. A. Lubitz, R. B. Schnabel, S. K. Agarwal, D. D. McManus, P. T. Ellinor, M. G. Larson, G. L. Burke, L. J. Launer, A. Hofman, D. Levy, J. S. Gottdiener, S. Kääb, D. Couper, T. B. Harris, E. Z. Soliman, B. H. Stricker, V. Gudnason, S. R. Heckbert, and E. J. Benjamin. "Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium." In: *Journal of the American Heart Association* 2.2 (2013), pp. 1–11. DOI: 10.1161/JAHA.112.000102.

[44] A. M. Chamberlain, S. K. Agarwal, A. R. Folsom, E. Z. Soliman, L. E. Chambless, R. Crow, M. Ambrose, and A. Alonso. "A clinical risk score for atrial fibrillation in a biracial prospective cohort (from the Atherosclerosis Risk in Communities [ARIC] Study)." In: *American Journal of Cardiology* 107.1 (2011), pp. 85–91. DOI: 10.1016/j.amjcard.2010.08.049.

[45] M. Yadava, A. B. Hughey, and T. C. Crawford. "Postoperative atrial fibrillation: Incidence, mechanisms, and clinical correlates." In: *Cardiology Clinics* 32.4 (2014), pp. 627–636. DOI: 10.1016/j.ccl.2014.07.002.

[46] C. R. Asher, D. P. Miller, R. a. Grimm, D. M. Cosgrove, and M. K. Chung. "Analysis of risk factors for development of atrial fibrillation early after cardiac valvular surgery." In: *Am J Cardiol* 82 (1998), pp. 892–895. DOI: 10.1016/S0002-9149(98)00498-6.

[47] D. Brathwaite and C. Weissman. "The new onset of atrial arrhythmias following major noncardiothoracic surgery is associated with increased mortality." In: *Chest* 114.2 (1998), pp. 462–468. DOI: 10.1378/chest.114.2.462.

[48] K. K. Christians, B Wu, E. J. Quebbeman, and K. J. Brasel. "Postoperative atrial fibrillation in noncardiothoracic surgical patients." In: *American Journal of Surgery* 182.6 (2001), pp. 713–5. DOI: 10.1016/S0002-9610(01)00799-1.

[49] S. R. Walsh, J. E. Oates, J. A. Anderson, S. D. Blair, C. A. Makin, and C. J. Walsh. "Postoperative arrhythmias in colorectal surgical patients: Incidence and clinical correlates." In: *Colorectal Disease* 8.3 (2006), pp. 212–216. DOI: 10.1111/j.1463-1318.2005.00881.x.

[50] S. E. Mayson, A. J. Greenspon, S. Adams, M. V. DeCaro, M. Sheth, H. H. Weitz, and D. J. Whellan. "The changing face of postoperative atrial fibrillation prevention: A review of current medical therapy." In: *Cardiology in Review* 15.5 (2007), pp. 231–241. DOI: 10.1097/CRD.0b013e31813e62bb.

[51] J. P. Mathew, R. Parks, J. S. Savino, A. S. Friedman, C. Koch, D. T. Mangano, and W. S. Browner. "Atrial Fibrillation Following Coronary Artery Bypass Graft Surgery." In: *JAMA* 276.4 (1996), p. 300. DOI: 10.1001/jama.1996.03540040044031.

[52] C. Chelazzi, G. Villa, and A. R. De Gaudio. "Postoperative Atrial Fibrillation." In: *ISRN Cardiology* 2011 (2011), pp. 1–10. DOI: 10.5402/2011/203179.

[53] J. L. Cox. "A perspective of postoperative atrial fibrillation in cardiac operations." In: *The Annals of Thoracic Surgery* 56.3 (1993), pp. 405–409. DOI: 10.1016/0003-4975(93)90871-E.

[54] J. P. Tsikouris, J. Kluger, J. Song, and C. M. White. "Changes in P-wave dispersion and P-wave duration after open heart surgery are associated with the peak incidence of atrial fibrillation." In: *Heart and Lung: Journal of Acute and Critical Care* 30.6 (2001), pp. 466–471. DOI: 10.1067/mhl.2001.118363.

[55] P. Kirchhof, S. Benussi, D. Kotecha, A. Ahlsson, D. Atar, B. Casadei, M. Castella, H. C. Diener, H. Heidbuchel, J. Hendriks, G. Hindricks, A. S. Manolis, J. Oldgren, B. A. Popescu, U. Schotten, B. Van Putte, P. Vardas, S. Agewall, J. Camm, G. B. Esquivias, W. Budts, S. Carerj, F. Casselman, A. Coca, R. De Caterina, S. Deftereos, D. Dobrev, J. M. Ferro, G. Filippatos, D. Fitzsimons, B. Gorenek, M. Guenoun, S. H. Hohnloser, P. Kolh, G. Y. Lip, A. Manolis, J. McMurray, P. Ponikowski, R. Rosenhek, F. Ruschitzka, I. Savelieva, S. Sharma, P. Suwalski, J. L. Tamargo, C. J. Taylor, I. C. Van Gelder, A. A. Voors, S. Windecker, J. L. Zamorano, and K. Zeppenfeld. "2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS." In: *European Heart Journal* 37.38 (2016), pp. 2893–2962. DOI: 10.1093/eurheartj/ehw210.

[56] S. J. Connolly, I. Cybulsky, A. Lamy, R. S. Roberts, B. O'Brien, S. Carroll, E. Crystal, K. E. Thorpe, and M. Gent. "Double-blind, placebo-controlled, randomized trial of prophylactic metoprolol for reduction of hospital length of stay after heart surgery: The β-Blocker Length Of Stay (BLOS) study." In: *American Heart Journal* 145.2 (2003), pp. 226–232. ISSN: 00028703. DOI: 10.1067/mhj.2003.147.

[57] P. D. Bryson. *Comprehensive Reviews in Toxicology for Emergency Clinicians*. 3rd ed. Washington, DC: Taylor & Francis, 1997, p. 167.

[58] L. S. Goodman and L. L. Brunton. *Goodman & Gilman's Pharmacological Basis of Therapeutics*. Ed. by L. L. Brunton, B. A. Chabner, and B. C. Knollmann. 12th ed. New York: McGraw-Hill, 2011, p. 2084.

[59] B. G. Katzung and A. J. Trevor. *Basic & Clinical Pharmacology*. Ed. by B. G. Katzung and A. J. Trevor. 13th ed. New York: McGraw-Hill Medical, 2015, p. 1203.

[60] P. Polster and J. Broekhuysen. "The adrenergic antagonism of amiodarone." In: *Biochemical Pharmacology* 25.2 (1976), pp. 131–134. DOI: 10.1016/0006-2952(76)90279-3.

[61] C. T. January, L. S. Wann, J. S. Alpert, H. Calkins, J. E. Cigarroa, J. C. Cleveland, J. B. Conti, P. T. Ellinor, M. D. Ezekowitz, M. E. Field, K. T. Murray, R. L. Sacco, W. G. Stevenson, P. J. Tchou, C. M. Tracy, J. L. Anderson, J. L. Halperin, N. M. Albert, B. Bozkurt, R. G. Brindis, M. A. Creager, L. H. Curtis, D. DeMets, R. A. Guyton, J. S. Hochman, R. J. Kovacs, E. M. Ohman, S. J. Pressler, F. W. Sellke, W. K. Shen, and C. W. Yancy. *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*. Vol. 130. 23. 2014, e199–e267. DOI: 10.1161/CIR.0000000000000041.

[62] R. Latini, G. Tognoni, and R. E. Kates. "Clinical Pharmacokinetics of Amiodarone." In: *Clinical Pharmacokinetics* 9.2 (1984), pp. 136–156. DOI: 10.2165/00003088-198409020-00002.

[63] F. Andreasen, H. Agerbaek, P. Bjerregaard, and H. Gøtzsche. "Pharmacokinetics of amiodarone after intravenous and oral administration." In: *European Journal of Clinical Pharmacology* 19 (1981), pp. 293–299. DOI: 10.1007/BF00562807.

[64] M. D. Freedman and J. C. Somberg. "Pharmacology and Pharmacokinetics of Amiodarone." In: *The Journal of Clinical Pharmacology* 31.11 (1991), pp. 1061–1069. DOI: 10.1002/j.1552-4604.1991.tb03673.x.

[65] Drugs.com. *Amiodarone information from Drugs.com*. URL: https://www.drugs.com/ppa/amiodarone.html (visited on 09/25/2018).

[66] G. Frendl, A. C. Sodickson, M. K. Chung, A. L. Waldo, B. J. Gersh, J. E. Tisdale, H. Calkins, S. Aranki, T. Kaneko, S. Cassivi, S. C. Smith, D. Darbar, J. O. Wee, T. K. Waddell, D. Amar, and D. Adler. "2014 AATS guidelines for the prevention and management of perioperative atrial fibrillation and flutter for thoracic surgical procedures." In: *Journal of Thoracic and Cardiovascular Surgery* 148.3 (2014), e153–e193. DOI: 10.1016/j.jtcvs.2014.06.036.

[67] P. Meurin, H. Weber, N. Renaud, F. Larrazet, J. Y. Tabet, P. Demolis, and A. B. Driss. "Evolution of the postoperative pericardial effusion after day 15: The problem of the late tamponade." In: *Chest* 125.6 (2004), pp. 2182–2187. DOI: 10.1378/chest.125.6.2182.

[68] R. B. Schnabel, L. M. Sullivan, D. Levy, M. J. Pencina, J. M. Massaro, R. B. D'Agostino, C. Newton-Cheh, J. F. Yamamoto, J. W. Magnani, T. M. Tadros, W. B. Kannel, T. J. Wang, P. T. Ellinor, P. A. Wolf, R. S. Vasan, and E. J. Benjamin. "Development of a risk score for atrial fibrillation in the community; The Framingham Heart Study." In: *Lancet* 373.9665 (2009), pp. 739–745. DOI: 10.1016/S0140-6736(09)60443-8.Development.

[69] R. B. Schnabel, T. Aspelund, G. Li, L. M. Sullivan, A. Suchy-Dicey, T. B. Harris, M. J. Pencina, R. B. D'Agostino, D. Levy, W. B. Kannel, T. J. Wang, R. A. Kronmal, P. A. Wolf, G. L. Burke, L. J. Launer, R. S. Vasan, B. M. Psaty, E. J. Benjamin, V. Gudnason, and S. R. Heckbert. "Validation of an atrial fibrillation risk algorithm in Whites and African Americans." In: *Archives of Internal Medicine* 170.21 (2010), pp. 1909–1917. DOI: 10.1001/archinternmed.2010.434.

[70] D. Amar, W. Shi, C. W. Hogue, H. Zhang, R. S. Passman, B. Thomas, P. B. Bach, R. Damiano, and H. T. Thaler. "Clinical prediction rule for atrial fibrillation after coronary artery bypass grafting." In: *Journal of the American College of Cardiology* 44.6 (2004), pp. 1248–1253. DOI: 10.1016/j.jacc.2004.05.078.

[71] S.-K. Chua, K.-G. Shyu, M.-J. Lu, L.-M. Lien, C.-H. Lin, H.-H. Chao, and H.-M. Lo. "Clinical utility of CHADS2 and CHA2DS2-VASc scoring systems for predicting postoperative atrial fibrillation after cardiac surgery." In: *The Journal of Thoracic and Cardiovascular Surgery* 146.4 (2013), 919–926.e1. DOI: 10.1016/j.jtcvs.2013.03.040.

[72] S. Helgadottir, M. I. Sigurdsson, I. L. Ingvarsdottir, D. O. Amar, and T. Gudbjartsson. "Atrial fibrillation following cardiac surgery: risk analysis and long-term survival." In: *Journal of Cardiothoracic Surgery* 7.1 (2012), pp. 87–93. DOI: 10.1186/1749-8090-7-87.

[73] G. Mariscalco, F. Biancari, M. Zanobini, M. Cottini, G. Piffaretti, M. Saccocci, M. Banach, C. Beghi, and G. D. Angelini. "Bedside tool for predicting the risk of postoperative atrial fibrillation after cardiac surgery: the POAF score." In: *Journal*

of the American Heart Association 3.2 (2014), pp. 1–10. DOI: 10.1161/JAHA.113.000752.

[74] M. Fukunami, T. Yamada, M. Ohmori, K. Kumagai, K. Umemoto, A. Sakai, N. Kondoh, T. Minamino, and N. Hoki. "Detection of patients at risk for paroxysmal atrial fibrillation during sinus rhythm by P wave-triggered signal-averaged electrocardiogram." In: *Circulation* 83.1 (1991), pp. 162–169. DOI: 10.1161/01.CIR.83.1.162.

[75] M. Klein, S. J. L. Evans, S. Blumberg, L. Cataldo, and M. M. Bodenheimer. "Use of P-wave-triggered, P-wave signal-averaged electrocardiogram to predict atrial fibrillation after coronary artery bypass surgery." In: *American Heart Journal* 129.5 (1995), pp. 895–901. DOI: 10.1016/0002-8703(95)90109-4.

[76] P. J. Stafford, S Kolvekar, J Cooper, J Fothergill, F Schlindwein, D. P. DeBono, T. J. Spyt, and C. J. Garratt. "Signal averaged P wave compared with standard electrocardiography or echocardiography for prediction of atrial fibrillation after coronary bypass grafting." In: *Heart (British Cardiac Society)* 77.5 (1997), pp. 417–22. DOI: 10.1136/hrt.77.5.417.

[77] J. S. Steinberg, S. Zelenkofske, S. C. Wong, M. Gelernt, R. Sciacca, and E. Menchavez. "Value of the P-wave signal-averaged ECG for predicting atrial fibrillation after cardiac surgery." In: *Circulation* 88.6 (1993), pp. 2618–2622. DOI: 10.1161/01.CIR.88.6.2618.

[78] G. K. Andrikopoulos, P. E. Dilaveris, D. J. Richter, E. J. Gialafos, a. G. Synetos, and J. E. Gialafos. "Increased variance of P wave duration on the electrocardiogram distinguishes patients with idiopathic paroxysmal atrial fibrillation." In: *Pacing and Clinical Electrophysiology* 23.7 (2000), pp. 1127–1132. DOI: 10.1111/j.1540-8159.2000.tb00913.x.

[79] K Aytemir, N Ozer, E Atalar, E Sade, S Aksoyek, K Ovünç, A Oto, F Ozmen, and S Kes. "P wave dispersion on 12-lead electrocardiography in patients with paroxysmal atrial fibrillation." In: *Pacing and Clinical Electrophysiology: PACE* 23 (2000), pp. 1109–1112. DOI: 10.1111/j.1540-8159.2000.tb00910.x.

[80] A. E. Buxton and M. E. Josephson. "The role of P wave duration as a predictor of postoperative atrial arrhythmias." In: *Chest* 80.1 (1981), pp. 68–73. DOI: 10.1378/chest.80.1.68.

[81] J Chandy, T Nakai, R. J. Lee, W. H. Bellows, S Dzankic, and J. M. Leung. "Increases in P-wave dispersion predict postoperative atrial fibrillation after coronary artery bypass graft surgery." In: *Anesth Analg* 98.2 (2004), pp. 303–10. DOI: 10.1213/01.ANE.0000096195.47734.2F.

[82] C.-M. Chang, S.-H. Lee, M.-J. Lu, C.-H. Lin, H.-H. Chao, J.-J. Cheng, P. Kuan, and C.-R. Hung. "The role of P wave in prediction of atrial fibrillation after coronary artery surgery." In: *International Journal of Cardiology* 68.3 (1999), pp. 303–308. DOI: 10.1016/S0167-5273(98)00301-5.

[83] P. E. Dilaveris, E. J. Gialafos, S. K. Sideris, A. M. Theopistou, G. K. Andrikopoulos, M. Kyriakidis, J. E. Gialafos, and P. K. Toutouzas. "Simple electrocardiographic markers for the prediction of paroxysmal idiopathic atrial fibrillation." In: *American Heart Journal* 135.5 (1998), pp. 733–738. DOI: 10.1016/S0002-8703(98)70030-4.

[84] T Hiraki, H Ikeda, M Ohga, T Hamada, I Kubara, T Yoshida, H Ajisaka, A Tanabe, M Kanahara, and T Imaizumi. "Frequency- and time-domain analysis of P wave in patients with paroxysmal atrial fibrillation." In: *PACE: Pacing and Clinical Electrophysiology* 21.1 Pt 1 (1998), pp. 56–64.

[85] R. Passman, J. Beshai, B. Pavri, and S. Kimmel. "Predicting post-coronary bypass surgery atrial arrhythmias from the preoperative electrocardiogram." In: *American Heart Journal* 142.5 (2001), pp. 806–810. DOI: 10.1067/mhj.2001.118736.

[86] P. Stafford, P. Denbigh, and R. Vincent. "Frequency analysis of the P wave: Comparative techniques." In: *PACE: Pacing and Clinical Electrophysiology* 18.2 (1995), pp. 261–270. DOI: 10.1111/j.1540-8159.1995.tb02516.x.

[87] V. Vassilikos, G. Dakos, I. Chouvarda, L. Karagounis, H. Karvounis, N. Maglaveras, S. Mochilas, P. Spanos, and G. Louridas. "Can P wave wavelet analysis predict atrial fibrillation after coronary artery bypass grafting?" In: *PACE: Pacing and Clinical Electrophysiology* 26.1 pt. 2 (2003), pp. 305–309. DOI: 10.1007/s11748-013-0272-y.

[88] D. Chamchad, J. C. Horrow, L. E. Samuels, and L. Nakhamchik. "Heart rate variability measures poorly predict atrial fibrillation after off-pump coronary artery bypass grafting." In: *Journal of Clinical Anesthesia* 23.6 (2011), pp. 451–455. DOI: 10.1016/j.jclinane.2010.12.016.

[89] T. F.o.t.E.S.o.C.t.N.A.S.o. P. Electrophysiology. "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use." In: *European Heart Journal* 17 (1996), pp. 354–381. DOI: 10.1161/01.CIR.93.5.1043.

[90] G. B. Moody, A. L. Goldberger, S. McClennen, and S. P. Swiryn. "Predicting the onset of paroxysmal atrial fibrillation: The computers in cardiology challenge 2001." In: *Computers in Cardiology* 28 (2001), pp. 113–116. DOI: 10.1109/CIC.2001.977604.

[91] B Hickey and C Heneghan. "Screening for paroxysmal atrial fibrillation using atrial premature contractions and spectral measures." In: *Computers in Cardiology 2002* 29.1 (2002), pp. 217–220.

[92] N. Montazeri, M. B. Shamsollahi, G. Carrault, and A. I. Hernández. "Paroxysmal atrial fibrillation prediction using Kalman Filter." In: *Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies*. 2011, pp. 1–5. DOI: 10.1145/2093698.2093787.

[93] G. Schreier, P. Kastner, and W. Marko. "An Automatic ECG Processing Algorithm to Identify Patients Prone to Paroxymal Atrial Fibrillation." In: *Computers in Cardiology 2001* 28 (2001), pp. 133–135. DOI: 10.1109/CIC.2001.977609.

[94] T. Thong, J. McNames, M. Aboy, and B. Goldstein. "Prediction of Paroxysmal Atrial Fibrillation by Analysis of Atrial Premature Complexes." In: *Computers in Cardiology 2001* 51.4 (2001), pp. 561–569.

[95] W. Zong, R. Mukkamala, and R. Mark. "A methodology for predicting paroxysmal atrial fibrillation based on ECG arrhythmia feature analysis." In: *Computers in Cardiology 2001* 28.1 (2001), pp. 125–128. DOI: 10.1109/CIC.2001.977607.

[96] H. Costin, C. Rotariu, and A. Pasarica. "Atrial fibrillation onset prediction using variability of ECG signals." In: *The 8th International Symposium on Advanced Topics in Electrical Engineering*. 2013, pp. 1–4. DOI: 10.1109/ATEE.2013.6563419.

[97] K. Lynn and H. Chiang. "A Two Stage Solution Algorithm for Paroxymal Atrial Fibrillation Prediction." In: *Computers in Cardiology 2001* 28 (2001), pp. 405–407. DOI: 10.1109/CIC.2001.977678.

[98] C. Maier, M. Bauch, and H. Dickhaus. "Screening and prediction of paroxysmal atrial fibrillation by analysis of heart rate variability parameters." In: *Computers in Cardiology 2001* 28 (2001), pp. 129–132. DOI: 10.1109/CIC.2001.977608.

[99] Y. H. Yang ACC. "Prediction of paroxymal atrial fibrillation by footprint analysis." In: *Computers in Cardiology 2001* 28 (2001), pp. 401–404. DOI: 10.1109/CIC.2001.977677.

[100] P. S. Hamilton and W. J. Tompkins. "Quantitative Investigation of QRS Detection Rules Using the MIT/BIH Arrhythmia Database." In: *IEEE Transactions on Biomedical Engineering* BME-33.12 (1986), pp. 1157–1165. DOI: 10.1109/TBME.1986.325695.

[101] J. Pan and W. J. Tompkins. "A real-time QRS detection algorithm." In: *IEEE Transactions on Biomedical Engineering* 3.3 (1985), pp. 230–236. DOI: 10.1109/TBME.1985.325532.

[102] C Li, C Zheng, and C Tai. "Detection of ECG characteristic points using wavelet transforms." In: *IEEE Trans Biomed Eng* 42.Bmei (1995), pp. 21–28. DOI: 10.1109/10.362922. URL: 7851927.

[103] J. P. Martínez, R. Almeida, S. Olmos, A. P. Rocha, and P. Laguna. "A Wavelet-Based ECG Delineator Evaluation on Standard Databases." In: *IEEE Transactions on Biomedical Engineering* 51.4 (2004), pp. 570–581. DOI: 10.1109/TBME.2003.821031.

[104] A. Martínez, R. Alcaraz, and J. J. Rieta. "Application of the phasor transform for automatic delineation of single-lead ECG fiducial points." In: *Physiological Measurement* 31.11 (2010), pp. 1467–1485. DOI: 10.1088/0967-3334/31/11/005.

[105] A. Martínez, R. Alcaraz, and J. J. Rieta. "Gaussian modeling of the P-wave morphology time course applied to anticipate paroxysmal atrial fibrillation." In: *Computer Methods in Biomechanics and Biomedical Engineering* 18.16 (2015), pp. 1775–1784. DOI: 10.1080/10255842.2014.964219.

[106] M. Ovreiu, M. Petre, D. Simon, D. Sessler, and C. Bashour. "Prediction of Paroxysmal Atrial Fibrillation Onset in Postoperative Patients Using Neuro-Fuzzy Modeling." In: *Second AMA-IEEE Medical Technology Conference on Delivering on the Promise of Cost Effective Quality Healthcare*. 2011, p. 1.

[107] A. Martínez, D. Abásolo, R. Alcaraz, and J. J. Rieta. "Alteration of the P-wave non-linear dynamics near the onset of paroxysmal atrial fibrillation." In: *Medical Engineering and Physics* 37.7 (2015), pp. 692–697. DOI: 10.1016/j.medengphy.2015.03.021.

[108] R. Alcaraz, A. Martínez, and J. J. Rieta. "Role of the P-wave high frequency energy and duration as noninvasive cardiovascular predictors of paroxysmal atrial fibrillation." In: *Computer Methods and Programs in Biomedicine* 119.2 (2015), pp. 110–119. DOI: 10.1016/j.cmpb.2015.01.006.

[109] R. Alcaraz, A. Martínez, and J. J. Rieta. "The P Wave Time-Frequency Variability Reflects Atrial Conduction Defects before Paroxysmal Atrial Fibrillation." In: *Annals of Noninvasive Electrocardiology* 20.5 (2015), pp. 433–445. DOI: 10.1111/anec.12240.

[110] G. Peretto, A. Durante, L. Limite, and D. Cianflone. "Postoperative arrhythmias after cardiac surgery: Incidence, risk factors, and therapeutic management." In: *Cardiology Research and Practice* 2014.i (2014). DOI: 10.1155/2014/615987.

[111] C. Orphanidou. *Signal Quality Assessment in Physiological Monitoring: State of the Art and Practical Considerations*. 1st ed. Cham: Springer International Publishing, 2018, p. 63. ISBN: 978-3-319-68414-7. DOI: 10.1007/978-3-319-68415-4.

[112] F. Censi, C. Ricci, G. Calcagnini, M. Triventi, R. P. Ricci, M. Santini, and P. Bartolini. "Time-Domain and Morphological Analysis of the P-Wave. Part I : Technical Aspects for Automatic Quantification of P-Wave Features." In: *PACE: Pacing and Clinical Electrophysiology* 31.7 (2008), pp. 874–883. DOI: 10.1111/j.1540-8159.2008.01102.x.

[113] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-k. Peng, and H. E. Stanley. "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals." In: *Circulation* 101.23 (2000), e215–e220.

[114] C. Orphanidou, T. Bonnici, P. Charlton, D. Clifton, and D. Vallance. "Signal Quality Indices for the Electrocardiogram and Photoplethysmogram : Derivation and Applications to Wireless Monitoring." In: *IEEE Journal of Biomedical and Health Informatics* 19.3 (2015), pp. 832–838. DOI: 10.1109/JBHI.2014.2338351.

[115] D. Belo and H. Gamboa. "Noise detection on ECG based on agglomerative clustering of morphological features." In: *Computers in Biology and Medicine* 87 (2017), pp. 322–334. DOI: 10.1016/j.compbiomed.2017.06.009.

[116] G. E.A.P. A. Batista, R. C. Prati, and M. C. Monard. "A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data." In: *Sigkdd Explorations* 6.1 (2004), pp. 20–29.

[117] N. V. Chawla. "Data Mining for Imbalanced Datasets : An Overview." In: *Maimon O., Rokach L. (eds) Data Mining and Knowledge Discovery Handbook. Springer, Boston, MA*. 2005, pp. 875–886. DOI: 10.1007/978-0-387-09823-4.

[118] H. He and E. A. Garcia. "Learning from Imbalanced Data." In: *IEEE Transactions on Knowledge and Data Engineering* 21.9 (2009), pp. 1263–1284. DOI: 10.1109/TKDE.2008.239.

[119] P. H. Charlton. "Continuous Respiratory Rate Monitoring to Detect Clinical Deteriorations using Wearable Sensors." Doctoral dissertation. King's College London, 2017, pp. 1–239.

[120] D. Filos, I. Chouvarda, D. Tachmatzidis, V. Vassilikos, and N. Maglaveras. "Beat-to-beat P-wave morphology as a predictor of paroxysmal atrial fibrillation." In: *Computer Methods and Programs in Biomedicine* 151 (2017), pp. 111–121. DOI: 10.1016/j.cmpb.2017.08.016.

[121] C. B. de Vos, R. Pisters, R. Nieuwlaat, M. H. Prins, R. G. Tieleman, R. J. S. Coelen, A. C. van den Heijkant, M. A. Allessie, and H. J. Crijns. "Progression From Paroxysmal to Persistent Atrial Fibrillation: Clinical Correlates and Prognosis." In: *Journal of the American College of Cardiology* 55.8 (2010), pp. 725–731. DOI: 10.1016/j.jacc.2009.11.040.

[122] P. Laguna, R. Mark, A. Goldberg, and G. Moody. "A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG." In: *Computers in Cardiology 1997* 24 (1997), pp. 673–676. DOI: 10.1109/CIC.1997.648140.

[123] S Sovilj, G Rajsman, and R Magjarevic. "Multiparameter Prediction Model for Atrial Fibrillation after CABG." In: *Computers in Cardiology, 2006*. 2006, pp. 489–492.

[124] G. Conte, A. Luca, S. Yazdani, M. L. Caputo, F. Regoli, T. Moccetti, L. Kappenberger, J. M. Vesin, and A. Auricchio. "Usefulness of P-Wave Duration and Morphologic Variability to Identify Patients Prone to Paroxysmal Atrial Fibrillation." In: *American Journal of Cardiology* 119.2 (2017), pp. 275–279. DOI: 10.1016/j.amjcard.2016.09.043.

[125] H. Sakoe and S. Chiba. "Dynamic Programming Algorithm Optimization for Spoken Word Recognition." In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 26.1 (1978), pp. 43–49. DOI: 10.1109/TASSP.1978.1163055.

[126] K. K. Paliwal, A. Agarwal, and S. S. Sinha. "A Modification over Sakoe and Chiba's Dynamic Time Warping Algorithm for Isolated Word Recognition." In: *Signal Processing* 4 (1982), pp. 1259–1261. DOI: 10.1016/0165-1684(82)90009-3.

[127] R. Alcaraz, A. Martínez, and J. J. Rieta. "The lagged central tendency measure applied to assess P-wave duration variability improves paroxysmal atrial fibrillation onset prediction." In: *Computing in Cardiology* 42 (2015), pp. 493–496. DOI: 10.1109/CIC.2015.7410955.

[128] S. Holm. "A Simple Sequentially Rejective Multiple Test Procedure." In: *Scandinavian Journal of Statistics* 6.6 (1979), pp. 65–70. DOI: 10.2307/4615733.

[129] Z. Šidák. "Rectangular confidence regions for the means of multivariate normal distributions." In: *Journal of the American Statistical Association* 62.318 (1967), pp. 626–533. DOI: 10.2307/2283989.

[130] D. Tecelão, A. Mendes, D. Martins, C. Fu, C. A. Chaddock, M. M. Picchioni, C. McDonald, S. Kalidindi, R. Murray, and D. P. Prata. "The effect of psychosis associated CACNA1C, and its epistasis with ZNF804A, on brain function." In: *Genes, Brain and Behavior* September (2018), e12510. DOI: 10.1111/gbb.12510.

[131] E. N. Prystowsky. "Management of atrial fibrillation: therapeutic options and clinical decisions." In: *The American Journal of Cardiology* 85.10 (2000), pp. 3–11. DOI: 10.1016/S0002-9149(00)00908-5.

[132] P. G. Platonov. "P-wave morphology: Underlying mechanisms and clinical implications." In: *Annals of Noninvasive Electrocardiology* 17.3 (2012), pp. 161–169. DOI: 10.1111/j.1542-474X.2012.00534.x.

[133] P. E. Dilaveris and J. E. Gialafos. "Future concepts in P wave morphological analyses." In: *Cardiac Electrophysiology Review* 6.3 (2002), pp. 221–224. DOI: 10.1023/A:1016320807103.

[134] K. C. Bilchick and R. D. Berger. "Heart rate variability." In: *Journal of Cardiovascular Electrophysiology* 17.6 (2006), pp. 691–694. DOI: 10.1111/j.1540-8167.2006.00501.x.

[135] P.-S. Chen, L. S. Chen, M. C. Fishbein, S.-F. Lin, and S. Nattel. "Role of the autonomic nervous system in atrial fibrillation: pathophysiology and therapy." In: *Circ Res* 114.9 (2014), pp. 1500–1515. DOI: 10.1161/CIRCRESAHA.114.303772.Role.

[136] S Vikman, T. Makikallio, S Yli-Mayry, S Pikkujamsa, A. Koivisto, P Reinikainen, K. Airaksinen, and H. Huikuri. "Altered complexity and correlation properties of RR interval dynamics before the spontaneous." In: *Circulation* 100 (1999), pp. 2079–2084.

[137] Y. V. Chesnokov. "Complexity and spectral analysis of the heart rate variability dynamics for distant prediction of paroxysmal atrial fibrillation with artificial intelligence methods." In: *Artificial Intelligence in Medicine* 43.2 (2008), pp. 151–165. DOI: 10.1016/j.artmed.2008.03.009.

[138] M. Mohebbi and H. Ghassemian. "Prediction of paroxysmal atrial fibrillation based on non-linear analysis and spectrum and bispectrum features of the heart rate variability signal." In: *Computer Methods and Programs in Biomedicine* 105.1 (2012), pp. 40–49. DOI: 10.1016/j.cmpb.2010.07.011.

[139] D. Amar, H. Zhang, S. Miodownik, and A. H. Kadish. "Competing autonomic mechanisms precede the onset of postoperative atrial fibrillation." In: *Journal of the American College of Cardiology* 42.7 (2003), pp. 1262–1268. DOI: 10.1016/S0735-1097(03)00955-0.

[140] Y. Xi and J. Cheng. "Dysfunction of the autonomic nervous system in atrial fibrillation." In: *Journal of Thoracic Disease* 7.2 (2015), pp. 193–198. DOI: 10.3978/j.issn.2072-1439.2015.01.12.

[141] D.-G. Shin, C.-S. Yoo, S.-H. Yi, J.-H. Bae, Y.-J. Kim, J.-S. Park, and G.-R. Hong. "Prediction of paroxysmal atrial fibrillation using nonlinear analysis of the R-R interval dynamics before the spontaneous onset of atrial fibrillation." In: *Circulation journal : Official Journal of the Japanese Circulation Society* 70.1 (2006), pp. 94–99. DOI: 10.1253/circj.70.94.

[142] A. E. Johnson, T. J. Pollard, L. Shen, L. W. H. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. Anthony Celi, and R. G. Mark. "MIMIC-III, a freely accessible critical care database." In: *Scientific Data* 3 (2016), pp. 1–9. DOI: 10.1038/sdata.2016.35.

[143] N Lakusic, D Mahovic, P Kruzliak, J Cerkez Habek, M Novak, and D Cerovec. "Changes in Heart Rate Variability after Coronary Artery Bypass Grafting and Clinical Importance of These Findings." In: *Biomed Res Int* 2015 (2015), pp. 1–7. DOI: 10.1155/2015/680515.

[144] C Dimmer, R Tavernier, N Gjorgov, G Van Nooten, D. L. Clement, and L Jordaens. "Variations of autonomic tone preceding onset of atrial fibrillation after coronary artery bypass grafting." In: *Am J Cardiol* 82.1 (1998), pp. 22–25. DOI: 10.1016/S0002-9149(98)00231-8.

[145] D. Chamchad, G. Djaiani, H. J. Jung, L. Nakhamchik, J. Carroll, and J. C. Horrow. "Nonlinear heart rate variability analysis may predict atrial fibrillation after coronary artery bypass grafting." In: *Anesthesia and Analgesia* 103.5 (2006), pp. 1109–1112. DOI: 10.1213/01.ane.0000239330.45658.76.

[146] K. Umetani, D. H. Singer, R. McCraty, and M. Atkinson. "Twenty-four hour time domain heart rate variability and heart rate: Relations to age and gender over nine decades." In: *Journal of the American College of Cardiology* 31.3 (1998), pp. 593–601. DOI: 10.1016/S0735-1097(97)00554-8.

[147] F. González, R. Alcaraz, and J. Rieta. "Electrocardiographic P-wave delineation based on adaptive slope Gaussian detection." In: *Computing in Cardiology* 44 (2017), pp. 1–4. DOI: 10.22489/CinC.2017.236-033.

[148] S. Nemati, M. M. Ghassemi, V. Ambai, N. Isakadze, O. Levantsevych, A. Shah, and G. D. Clifford. "Monitoring and detecting atrial fibrillation using wearable technology." In: *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2016), pp. 3394–3397. DOI: 10.1109/EMBC.2016.7591456.

[149] S. C. Tang, P. W. Huang, C. S. Hung, S. M. Shan, Y. H. Lin, J. S. Shieh, D. M. Lai, A. Y. Wu, and J. S. Jeng. "Identification of Atrial Fibrillation by Quantitative Analyses of Fingertip Photoplethysmogram." In: *Scientific Reports* 7.April (2017), pp. 1–7. DOI: 10.1038/srep45644.

[150] B. Paliakaitė, A. Petrėnas, J. Skibarkienė, T. Mickus, S. Daukantas, R. Kubilius, and V. Marozas. "Towards Long-term Monitoring of Atrial Fibrillation using Photoplethysmography." In: *Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2017)*. 2017, pp. 141–146. DOI: 10.5220/0006115601410146.

[151] J. M. Bland and D. G. Altman. "Agreement between methods of measurement with multiple observations per individual." In: *Journal of Biopharmaceutical Statistics* 17.4 (2007), pp. 571–582. DOI: 10.1080/10543400701329422.

[152] P. S. Myles and J. Cui. "Using the Bland-Altman method to measure agreement with repeated measures." In: *British Journal of Anaesthesia* 99.3 (2007), pp. 309–311. DOI: 10.1093/bja/aem214.

Appendix

Appendix A: Comparison of P-wave delineators

This appendix presents a comparison of two P-wave delineators: one using the phasor transform (PT), and other based on the wavelet transform. The aim of this study was to select the delineator which is best suited to be used in atrial fibrillation (AF) prediction analyses. The ability of the algorithms to delineate P-waves was compared using: 1) comparisons of P-wave delineations with manual annotations; and 2) visual inspection of P-wave delineations on electrocardiogram (ECG) recordings from healthy subjects and patients prone to AF.

A.1 Introduction

The P-wave from the ECG represents the electrical activity of the atria and provides information about atrial conduction. Recent studies have found that patients who develop AF have greater variability in P-wave morphology and time-intervals [25–27, 105, 107, 108]. Hence, the prediction of AF is highly reliant on the accurate delineation of P-waves. Measurements derived from wrongly delineated P-waves are often erroneous, which compromises the accuracy of the prediction methods. In addition, P-wave delineation must be performed automatically, so that AF might be predicted during unsupervised continuous monitoring.

Many different methods for delineating P-waves have been introduced. Their strategies range from the use of mathematical transforms, such as the PT [104] or the wavelet transform [103], to the use Gaussian models [147]. In this study we present a comparison of two P-wave delineators, one using the PT and other using the Dyadic wavelet transform (wavedet algorithm), with the aim of selecting the delineator which will be used in the AF prediction analyses of this thesis.

A.2 Methods

The ability of the PT and wavedet algorithms to delineate P-waves was compared using: 1) comparisons of the P-wave delineations provided by the algorithm with manual annotations from the QT database (QTDB) [122]; and 2) visual inspection of the P-wave delineations provided by the algorithms in the AF prediction database (AFPDB) [90].

A.2.1 Comparisons with manual annotations

The QTDB was used to compare the performance of the two delineators from real ECGs. The database was developed for wave limit validation purposes, and contains annotations by an expert cardiologist for at least 30 beats per recording, including P-wave onset, peak and offset [122].

The delineators were applied to the bandpass-filtered (between 0.5 and 40Hz) ECG signals and their accuracy to delineate P-wave onset, peak and offset was assessed using the location error (LE). LE was was computed as the subtraction of the automatic delineations (obtained with the algorithms) from the manual annotations (Figure A.1). Given that both delineators work on a single-channel basis, it has to remarked that the manual annotations on the QTDB were performed having the two available leads in sight. To minimise this discrepancy, the two channels were delineated individually and the LE was computed for each channel, but only the smaller of the two LEs was selected. In addition, the error in P-wave duration was also calculated.

Finally, the performance of the delineators was compared using the limits of agreement (LoA), a technique widely used to compare two measuring methods [151, 152]. The following statistics were computed:

1. Bias (corresponding to the mean of all LEs).
2. Standard deviation (SD) of LEs. Allows to compute the expected range of 95% of errors around the systematic bias, given by ±1.96SD ("2SD"). The standard deviation of errors is given by SD = $\sqrt{\text{total variance}}$, where the total variance is the sum of the within-subjects and between subjects variance. Hence, the standard deviation can be computed as:

$$\text{SD} = \sqrt{\text{within-subjects variance} + \text{between-subjects variance}} \tag{A.1}$$

The two variance components can be estimated using one-way analysis of variance:

$$\text{within-subjects variance} = MS_{residual} \quad , \quad \text{and} \tag{A.2}$$

$$\text{between-subjects variance} = \frac{MS_{subject} - MS_{residual}}{\frac{(\sum m_i)^2 - \sum m_i^2}{(n-1)\sum m_i}} \quad , \tag{A.3}$$

where $MS_{residual}$ is the mean square error, and $MS_{subject}$ is the difference between the mean squares for ECG records, n is the number of records, and m_i is the number of delineations for the i^{th} record.

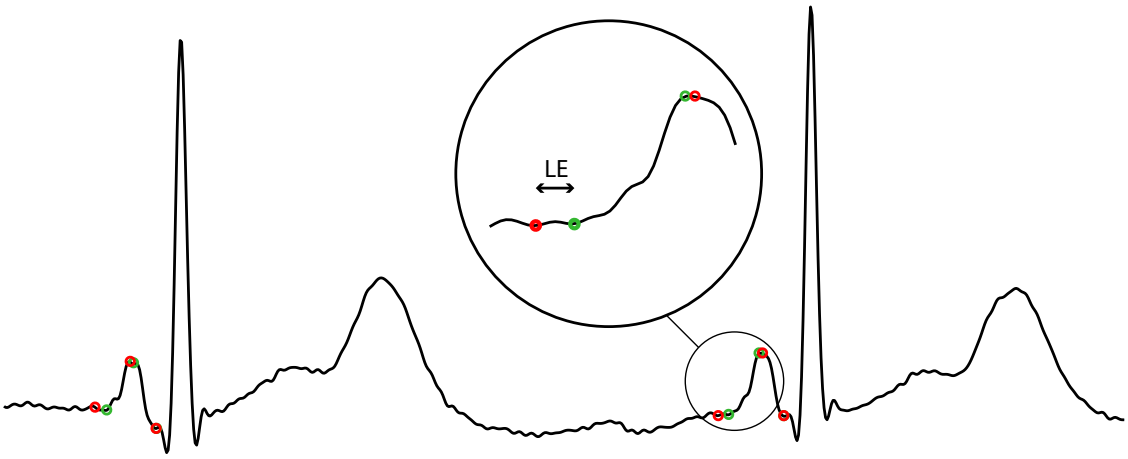

Figure A.1: Representative example of the calculation of the location error (LE), obtained by subtracting the delineation obtained using the automatic delineator (red) from the manual annotation (green).

A.2.2 Visual inspection of P-wave delineations

The AFPDB database contains ECG recordings from healthy subjects and patients who subsequently developed paroxysmal AF. Given that this analysis aimed to choose the best delineator to be used in AF prediction, it was important to inspect how the two delineators performed in that context. From the 61 records containing P-waves, we randomly selected 36, and visually inspected the delineations from both delineators. The number of records in which the overall P-wave delineations were bad was annotated. A record was considered to have bad P-wave delineations if any of the delineated fiducial points (onset, peak or offset) affected substantially the extraction of features (e.g., duration) for a large percentage of beats.

A.3 Results

Table A.1 shows the bias and SD of the P-wave delineation LEs for the PT and wavedet algorithms. The wavedet algorithm presented lower LE bias in all P-wave fiducial points, and lower SD in P-wave onset LE and duration.

Regarding the visual inspection of P-wave delineations in the AFPDB, the wavedet algorithm presented good P-wave delineations in 41.7% of records, while the delineations of the PT algorithm were good in 77.8%.

A.4 Discussion

During this analysis we aimed to select the delineator best suited to be applied in AF prediction analyses. From the LE analysis, we found that, even though the wavedet algorithm outperformed the PT algorithm, both approaches gave satisfactory results.

Table A.1: Comparison of the delineation performance of the phasor transform and wavedet algorithms, measured with the location error.

Phasor transform algorithm				
Statistics	**P-wave onset**	**P-wave peak**	**P-wave offset**	**P-wave duration**
Bias (ms)	-4.23	-4.90	-15.75	-2.87
SD (ms)	28.12	21.50	22.16	7.47
Wavedet algorithm				
Statistics	**P-wave onset**	**P-wave peak**	**P-wave offset**	**P-wave duration**
Bias (ms)	1.40	1.42	6.14	1.19
SD (ms)	26.84	25.17	23.68	7.12

However, the performance of the PT algorithm in the AFPDB was much better than that of the wavedet algorithm. Taken together, these results show that the PT algorithm is most suited to be applied in AF prediction analyses.