Characterizing the Risk of Atrial Fibrillation in Cardiac Patients with Exceptional Electrocardiogram Phenotypes

Lieke van den Biggelaar
l.a.j.v.d.biggelaar@tue.nl
Eindhoven University of Technology
Eindhoven, the Netherlands

Rianne M. Schouten
r.m.schouten@tue.nl
Eindhoven University of Technology
Eindhoven, the Netherlands

Ashley de Bie
ashley.d.bie@catharinaziekenhuis.nl
Catharina Hospital
Eindhoven, the Netherlands

R. Arthur Bouwman
arthur.bouwman@catharinaziekenhuis.nl
Catharina Hospital
Eindhoven, the Netherlands

Wouter Duivesteijn
w.duivesteijn@tue.nl
Eindhoven University of Technology
Eindhoven, the Netherlands

Abstract

We provide a transparent method to characterize Atrial Fibrillation (AF) caused by cardiac surgery, using Electrocardiogram (ECG) phenotypes. Current practice in the hospital is reactive rather than preventive and is based on a third party's proprietary alarms on vitals. Assistance with detection and prediction methods often lacks sufficient insights into their decisions toward the users, i.e., the hospital workers. This aspect is necessary to gain the trust of medical workers and patients in the decisions that are made. Our objective of transparently identifying risk factors for AF helps experts increase their understanding of the problem, and assists in decision-making about administering preventive medication to risk groups. With the deployment of the Exceptional Model Mining (EMM) framework on AF-related ECG phenotypes, we introduce a transparent and actionable method that assists the hospital in preventive treatment. We find several subgroups with EMM that align with known risk factors in the existing literature, confirming the ability of our method to identify risk groups of AF successfully. In addition, new hypotheses on found characteristics and combinations thereof have originated from the deployment. The hospital is advised to administer preventive medications to patients who match the descriptions of the risk groups found and perform follow-up clinical studies to validate the found hypotheses.

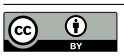
CCS Concepts

• **Information systems** → **Data mining**; *Association rules*; • **Computing methodologies** → Cluster analysis; Anomaly detection; Artificial intelligence.

Keywords

Atrial Fibrillation; Electrocardiogram Signals; Exceptional Model Mining; Signal Processing

ACM Reference Format:
Lieke van den Biggelaar, Rianne M. Schouten, Ashley de Bie, R. Arthur Bouwman, and Wouter Duivesteijn. 2025. Characterizing the Risk of Atrial Fibrillation in Cardiac Patients with Exceptional Electrocardiogram Phenotypes.



KDD '25, Toronto, ON, Canada

ACM ISBN 979-8-4007-1454-2/2025/08
https://doi.org/10.1145/3711896.3737200

In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2 (KDD '25), August 3–7, 2025, Toronto, ON, Canada.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3711896.3737200

KDD Availability Link:
The source code of this paper has been made publicly available at https://doi.org/10.5281/zenodo.15495124.

1 Introduction

Atrial Fibrillation (AF) is a cardiac arrhythmia that can be a side effect of cardiac surgery [1]. Up to 50% of cardiac patients develop AF within the first four weeks after surgery [44]. Once a patient has experienced AF, episodes can come and go for the remainder of their life. The resulting heart rhythm disturbance leads to turbulent blood flow, which can cause blood clots. A stroke is the main consequence of these blood clots [40]: 20% of AF patients get a stroke, making AF its leading cardiac cause [18]. AF episodes can be cured, but the risk of future episodes remains. Medication can prevent AF occurrence, thereby preventing risks introduced by AF [37]. The value of our method lies in identifying the risk groups of AF for which doctors can administer these preventive medications.

Current practice is reactive rather than preventive [37]. Monitoring systems are installed in the intensive care unit and operating rooms to record vitals and are equipped with alarms. These systems are created by a third party, using proprietary methods: alarms are triggered by variations in Electrocardiogram (ECG) signals whose properties are unknown to the users, i.e., the hospital workers. This method is sufficient to help patients who experience AF, but as the episode is already occurring, preventive medication cannot be administered. Predictive methods have been applied to assist doctors in dividing ECG features, and detecting and predicting AF [5, 20, 58]. This is possible due to the phenotypes in the ECG signals such as irregular Heart Rate Variability (HRV) and the absence of P-waves (more details to follow in Section 2). These solutions generally fall under the global modeling paradigm within data mining: one model is trained to make a prediction for every patient in the current data set and every new patient that might still emerge later. This one-size-fits-all approach is common: we learn a model, striving for a high predictive accuracy that generalizes to unseen examples, and then deploy that model globally. It is well-known, however, that this is not necessarily the correct approach in medicine [13]. To build trust between the deployed methods and the users, i.e., the medical workers, and to comply with ethical considerations about

the sensitive healthcare application, we should ensure that patients receive the care that is correct for them. In a move towards stratified medicine [55] we develop and deploy a method that can identify subgroups of patients with a higher risk of developing AF during or after surgery, aimed at detecting combinations of risk factors. With this information, doctors can provide preventive medications which may avoid the need to resort to curative treatment.

A suitable method for stratified medical assistance is Exceptional Model Mining (EMM) [16, 35]. EMM aims to identify actionable subgroups that show exceptional behavior, such as phenotypes on the ECG that indicate AF. Existing EMM methods are commonly applied to tabular representations of the data. However, our research concerns time-varying data: time series in the form of ECG signals and their characteristics. We address in this paper the challenge of adapting existing EMM methods to fit our data type.

The deployment of our method in the Catharina hospital in Eindhoven, the Netherlands has resulted in the discovery of various risk groups that are defined in terms of patient characteristics and have exceptional morphologies in their ECG. These subgroups align with existing literature, thus confirming the validity of our approach. We find several surprising risk group characteristics and combinations thereof that are interesting to research as hypotheses in future work on AF risk factors. Our results assist the hospital workers in determining the need for preventive medication for patients described by the subgroups.

This research contributes to the field of AF analysis and EMM. Our contribution includes: 1) deployment of EMM in the hospital to characterize AF; 2) discovery of previously unknown potential risk groups for AF during and after surgery; 3) application of EMM on ECG signals (time series); 4) introduction of new quality measures based on ECG phenotypes; 5) evaluation of the validity of subgroups using additional medical characteristics.

2 Background and Related Work

The process of pumping blood through the body, the *cardiac cycle*, works with pressure variations activated by an electrochemical pulse [28]. Every individual has a unique rhythm of the cardiac cycle, the *heart rhythm*, suitable for their body type and lifestyle, adapting to daily activities by increasing or decreasing the frequency of cardiac cycles, the *heart rate*.

We monitor the heart rate by measuring its electrochemical activity, through an *Electrocardiogram* (ECG). Using up to twelve *leads*, we record several angles resulting in different morphologies of this activity [41]. The signals are complex waveforms of elements, typically modeled as a *PQRST-complex*; Figure 1 displays an example. This model imitates the heart in two cardiac cycles.

Atrial Fibrillation (AF) is a heart arrhythmia that is often triggered after cardiac surgery. It is caused by a dysfunction in the node that gives the electrochemical pulse to activate the cardiac cycle. AF patients receive this signal through pulses originating outside of the sinoatrial node. These rapid pulses cause the atria, the heart's upper chambers, to fibrillate [28]. AF can be identified by variations in the morphology of the ECG, the *phenotypes*. This is a nontrivial task. Within a single patient, the ECGs that are typical of AF episodes commonly display a substantial higher degree of irregularity than the ECGs that are typical of regular sinus rhythm

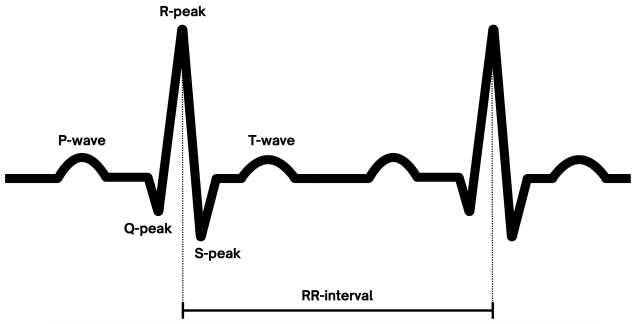


Figure 1: The ECG signal of two cardiac cycles showing each component in the PQRST-complex and the RR-interval.

heartbeats. However, simultaneously, ECGs differ drastically when compared across patients: what is a regular heartbeat for the one patient may not be a regular heartbeat for another patient. This diversity of signals between patients provides a challenge for extracting the AF-relevant aspects from ECGs. In the PQRST-complex of a patient's ECG, the P-waves represent the electrochemical signal: AF causes absence of this signal which causes absence of the P-wave in the ECG. As a consequence, a sequence of impulses can be generated in the heart that mimics the absent signal, causing fibrillatory waves (F-waves) to form between the T-wave and the QRS-complex [26]. The other visible phenotype is irregular Heart Rate Variability (HRV), shown by the RR-intervals. The duration of the RR-interval varies from the beginning to the end of an episode of AF, making it a consequence of AF rather than a precursor [51].

2.1 Predictive Methods

With new technologies, AF can be detected without continuous manual monitoring or third-party alarms. A simple method uses statistical analysis on the PQRST-complex in ECG signals [54]. Improvements include more complex frameworks based on the irregularities in the signals found using pattern analysis based on various phenotypes, such as RR-intervals [33] or missing P-waves [14]. These models can be based on autoregressive moving average models [33], hidden Markov models [14], and neural networks [11].

AF detection focuses on the phenotypes indicating AF, while prediction algorithms must find abnormalities in the ECG before AF. These algorithms use various characteristics in frequency, time, space, and other nonlinearities modeled using a support vector machine [10], random forest, multilayer perceptron, and k-nearest neighbor [50]. Many more applications of detection and prediction algorithms exist; a good survey can be found in [32]. The detection and prediction algorithms achieve great accuracy, but in a global, one-size-fits-all manner that befits data mining really well but may not be the best approach for medicine [13]. This complicates the deployment of these predictive methods in real-world practice.

2.2 Descriptive Methods

As opposed to the global modeling exemplified by predictive methods, another class of methods takes a local view of the data set, striving to describe only part of the data set at a time. This setting

Table 1: A toy example of the shape taken by our data set Ω, illustrated with two concrete patients. The data for each patient partitions into three modalities: the *descriptors* $a_1, \ldots, a_M$ describe Electronic Health Record data including medical history and demographics, the *targets* $\ell_1, \ldots, \ell_K$ are the heartbeats derived from the patient's ECG, and the third set $b_1, \ldots, b_L$ are binary AF complication occurrences at L time stages as annotated by medical staff or provided by a black-box algorithm. Exceptional Model Mining traverses the search space of candidate subgroups of patients *defined* in terms of the descriptors $a_1, \ldots, a_M$, and *evaluated* by a quality measure that rewards a subgroup for excelling in three components (cf. Section 5.1.4): 1) displaying exceptional Atrial Fibrillation phenotypes derived from the targets $\ell_1, \ldots, \ell_K$; 2) including a substantial number of patients; 3) including a substantial number of AF complications as represented by the binary complication occurrences $b_1, \ldots, b_L$.

patient	a_1=age	a_2=smokes?	…	a_M=blood loss	ℓ_1	ℓ_2	…	ℓ_K	b_1	b_2	…	b_L
p^1	30	no	…	high			…		0	0	…	0
p^2	46	yes	…	high			…		0	1	…	0
⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮

falls under the umbrella of *Local Pattern Mining* (LPM) [46]: inherently transparent methods leading to the discovery of interesting subgroups in the data set. *Frequent Itemset Mining* (FIM) [4] is such a transparent method that aims to discover combinations of items that occur frequently in the data. Such itemsets are turned into rules in *Association Rule Learning* (ARL) [4]: if all items on the left-hand-side (LHS) of the rule are present in a transaction, often, so are all items on the right-hand-side (RHS). FIM and ARL are *unsupervised* variants of LPM: all items play the same role in the algorithm and can in principle appear on both sides of the arrow. *Supervised* alternatives also exist in LPM: these partition the attributes of a data set in *descriptors* (LHS) and *targets* (RHS). *Subgroup Discovery* (SD) [6, 25, 31, 57] strives to find subgroups where a single target column has an exceptional distribution. *Exceptional Model Mining* (EMM) [16, 35] generalizes SD to a more complex target space: it seeks subgroups where some modeling over a target space displays exceptional behavior.

2.2.1 Application to Atrial Fibrillation. ARL is used to analyze stroke risks [29] and to detect cerebral infarction in patients with AF [38]. In the former, AF is included as a descriptor; the latter includes AF as a population requirement. An instance of SD similarly uses AF as a descriptor to find subgroups of patients with brain injury [20]. FIM is applied to identify risk factors for strokes in episodes of AF [39] and discover improvements in stroke preventive treatments after AF [23]. They aim to identify patients at risk for strokes out of a population of patients who have or had AF. Lastly, clustering is used to distinguish various types of AF experienced by patients [47, 56], finding various risk factors in the population of patients who all experienced AF.

3 Preliminaries: the Exceptional Model Mining Framework

Assume a data set Ω, which is a bag of N patients $p \in \Omega$ of the form $p = (a_1, ..., a_M, \ell_1, ..., \ell_K, b_1, ..., b_L)$. Here, $(a_1, ..., a_M)$ are the M *descriptors*, taking values from the domain $\mathcal{A} = \bigtimes_{m=1}^{M} \mathcal{A}_m$, where the product is a Cartesian product and each $\mathcal{A}_m$ can be any reasonable domain: integer values if a_m represents age, continuous values if a_m represent some sensor reading, categorical values if a_m represents smoking history, and so on. Collectively, $\mathcal{A}$ represents the Electronic Health Records (EHR). The attributes $(\ell_1, ..., \ell_K)$ are the heartbeats from the cardiac signal measured by the ECG, referred to as *targets*. Notice that partitioning a patient's ECG into K distinct heartbeats poses a nontrivial signal processing task; extensive details on how we approach this task can be found in Section 4. Lastly, binary AF complication occurrences at L stages $(b_1, ..., b_L)$ are used to evaluate subgroups. Specifics are indicated with superscript and subscript: patient j with attribute a_m^j about characteristic m. Table 1 illustrates the shape our data set takes for a small set of fictional example patients.

Candidate subgroups are generated by a guided search through a *description language* $\mathcal{D}$, encompassing several *descriptions* $D \in \mathcal{D}$. Mathematically, a description can be any function $D : \mathcal{A} \to \{0, 1\}$. In practice, $\mathcal{D}$ is typically limited to conjunctions of conditions on individual descriptors a_m: a description will take a form such as "age $\leq$ 25 AND smokes? = no". Descriptions, in turn, induce subgroups.

Definition 3.1 (Subgroup). The *subgroup* induced by a description D is the bag of patients $G_D \subseteq \Omega$ such that

$$G_D = \left(p^j \in \Omega \,\middle|\, D\left(a_1^j, ..., a_M^j\right) = 1 \right)$$

Candidate subgroups are evaluated through a *quality measure* φ.

Definition 3.2 (Quality Measure). A *quality measure* (QM) is a function φ: $\mathcal{D} \to \mathbb{R}$ assigning a numerical value to description D.

The QM quantifies the extent to which model behavior in the target space of subgroup G_D varies from that same behavior across the entire data set Ω. Subsequently, the goal of Exceptional Model Mining (EMM) is to report a list of exceptional subgroups (task definitions vary: top-q, top-q under diversity constraints, all subgroups passing a significance test, etcetera), discovered in a search through the space of candidate subgroups (strategies vary: Breadth First Search, Depth First Search, Beam Search, etcetera) guided by φ.

Data arrives at our doorstep in the form of a data set Ω which is not immediately ready for deployment of Exceptional Model Mining. EMM expects all data to be preprocessed in a flat-table form where every column (attribute) contains a single value for every row (observation), but the hospital delivers the target space $\ell_1, \ldots, \ell_K$ in the form of an ECG signal to be partitioned into K heartbeats.

The following two sections describe how we process this data and convert it into a form suitable for EMM: in Section 4, we convert the ECG signal into a flat-table data set as common in EMM, processing the heartbeats into numeric and binary columns representing heartbeat properties; in Section 5, we combine these properties into AF phenotypes, deriving quality measures that evaluate the exceptionality of AF phenotypes across subgroups of patients.

4 Cardiac Signal Preprocessing

Recorded signals suffer from distortions and faults that mask the underlying patterns. Noise in ECGs can be caused by power line interference, muscle artifact, baseline wander, and human error [30]. The leads can even be incorrectly secured, leading to inverted signals. This noise sabotages our analysis and needs to be handled while keeping the underlying signal intact. We apply several signal processing steps to overcome such noise.

First, we apply an altered version of NeuroKit ECG inversion [42, 43]: it checks if a waveform is inverted, and if so, inverts it back. Their backup check is removed for efficiency benefits.

The remaining signal noise can be categorized into *frequency noise* and *artifact noise*. Frequency noise obstructs patterns in heart rhythms, interfering with AF phenotypes. This operates at high frequencies: 50-60 Hz and over 100 Hz [48]. Artifact noise obstructs the small variations in the ECG signals by creating fluctuations around the baseline below 1 Hz [36]. Both can be removed by a filter that cuts off high or low frequencies, respectively.

We test several denoising methods. We implement discrete wavelet transform methods using wavelets with a curve similar to the PQRST-complex. The Daubechies3 [7] and Symlet3 [12] wavelets tend to overfit, leaving noise in the signal. The Biorthogonal4.4 [2] wavelet recreates sinus rhythms, which removes the irregular AF phenotypes in the ECG. Other filters include the Butterworth [8] band-pass filter which smooths the signal and lowers the peak amplitudes. Low-pass filtering tends to over-smooth the peaks leaving them undetectable, and high-pass filtering amplifies the peaks out of proportion, doing the same for peaks caused by noise. The Weighted Moving Average (WMA) [49] filter removes noise while keeping the form of the signal. A combination of Butterworth high-pass and Gaussian WMA filters proves to be the best. Butterworth high-pass eliminates baseline wander with a cutoff at 0.75; a value closest to the frequency of baseline wander, while keeping a buffer for weaker recordings. Gaussian WMA over ten instances with a sigma of 20 removes the remaining noise. These settings eliminate the noise in the signal while keeping its features close to the original. The resulting ECG signal contains clear and smooth heartbeats while keeping the AF phenotypes intact. Remaining recordings of under 30 seconds are removed: they risk having much noise compared to the number of features in the signal.

4.1 Distinguishing Arrhythmia Types

We identify ECG characteristics that separate AF episodes from sinus rhythms and other arrhythmias. These characteristics are irregular RR-intervals and P-wave absence (cf. Section 2).

A sinus rhythm features a heart rate of 60–100 bpm, PQ-intervals of 120–200 ms, QRS-complexes of 60–100 ms, and a P-wave duration of up to 120 ms [15]. Heart arrhythmias are identifiable by variations of these characteristics. Table 2 lists four arrhythmias with similar morphological abnormalities. Ventricular Fibrillation (VF) affects the heart's lower chambers; AF affects the upper chambers. In AF the P-wave is nonexistent; in VF the P-wave is shown to be independent of the QRS-complex (AV-dissociation). Patients with Atrial Flutter (AFl) or AF both experience shifts in their heartbeats, but AFl causes faster heartbeats with regular RR-intervals; these are highly irregular for AF. Left Bundle Branch Block (LBBB) is a slower left ventricle, shown by irregular RR-intervals and missing P-waves, similar to AF. The QRS-duration, T_{QRS}, is >120 ms for LBBB; it is <120 ms for AF.

Table 2: The ECG characteristics of the sinus rhythm and four heart arrhythmias with similar variations in the ECG: Atrial Fibrillation (AF), Ventricular Fibrillation (VF), Atrial Flutter (AFl), and Left Bundle Branch Block (LBBB).

Rhythm	RR-interval	P-wave	T_{QRS}
Sinus	regular	constant	60-100 ms
AF	irregular	non-existent	< 120 ms
VF	irregular	AV-dissociation	> 120 ms
AFl	regular	negative saw-tooth	< 120 ms
LBBB	irregular	non-existent	> 120 ms

4.2 QRS-complex

The QRS-complex is the main element of a heartbeat in the ECG signals and will function as the guideline for the detection of the P-waves of the heartbeat. We implement a NeuroKit method that uses local maxima to detect the R-peaks of each heartbeat [42, 43].

We compute the heart rate, RR_i, at heartbeat i as the distance between consecutive peaks, and denote the difference between consecutive intervals as ΔRR_i.

$$RR_i = R_i - R_{i-1}$$
$$\Delta RR_i = |RR_i - RR_{i-1}|$$

Additional rules exclude impossible heart rhythms and wrongfully selected R-peaks. This overcomes some of the challenges caused by the remaining noise and flatlines after signal processing. Most faulty RR-intervals come from selecting the wrong R-peak, causing too small intervals. Larger intervals can be caused by AF, but also by skipping beats and/or other diseases; since we can neither conclude nor exclude that such occurrences relate to AF, we choose to keep such intervals unaltered. According to domain experts, the RR-interval must be at least 400 ms. For smaller intervals, we compute the distance between both peaks and their previous or consecutive peak. We assume that the peak with the shortest distance is the faulty peak, and we remove it.

The Q- and S-peaks in the QRS-complex are necessary to identify the start and end of the QRS-complex, and the distance between a P-wave and its following QRS-complex. We incorporate an algorithm from NeuroKit based on the theory of a Wavelet-based ECG delineator [42, 43]. It does not select the Q- and S-peak of the first and the last QRS-complex in a recording. Since each recording includes many heartbeats, this exclusion of the two heartbeats at the far ends is an acceptable loss.

4.3 P-waves and F-waves

We limit the segment of one entire heartbeat to the interval between the end of the QRS-complex, S-peak, and the start of the next QRS-complex, Q-peak (SQ-interval). We check the validity of said SQ-intervals using rules: the S-peak must lie before the Q-peak, the next S-peak cannot lie before the Q-peak, and the SQ-interval duration must be realistic compared to the average SQ-interval of the patient: not longer than the average RR-interval. We add denoising filters to ease the identification of the waves from the remaining noise. We apply two Gaussian WMA filters: once over ten instances with a sigma of 20, and then again over 50 instances with a sigma of 25.

With the smooth signal, it is possible to apply a local maxima search technique. The segment is scaled to the range $[0, 1]$ using min-max scaling [24, Equation 2]. Each heartbeat is now handled similarly regardless of external factors that influence the strength and quality of the recording.

We compute the patient's average PQ-interval using the selected R-peaks. The hospital starts measuring ECGs 20 minutes before the surgery begins, giving us a baseline of clean heartbeats of each specific patient. To identify the average PQ-interval, we take the SQ-interval in these first 20 minutes and take the last wave in the range as the P-wave. All PQ-intervals in these first 20 minutes are averaged, thereby creating an average PQ-interval which we stretch a little to create the range in which a P-wave should be found. This is possible because the PQ-interval is not affected by AF. Ultimately, binary indicators on P-wave absence follow per heartbeat: P-wave existence (0), or P-wave absence (1). Three cases are considered:

- **P-wave existence (0):** A wave is detected in the PQ-range for the patient, and < 3 waves exist in the range.
- **F-wave occurrences (1):** A wave is detected in the PQ-range and ≥ 3 waves exist in the range.
- **P-wave absence (1):** No wave is detected in the PQ-range.

We include a minimum of three waves instead of one because in this range, other waves can exist that are regular, and replacement F-waves always occur in larger multitudes. In addition, we create a second binary list keeping track of the heartbeats showing F-waves. Of the three cases mentioned above, the P-wave absence case is also stored as (0) in this list as no F-waves occur.

5 Incorporating Atrial Fibrillation Phenotypes in Exceptional Model Mining

To characterize subgroups with exceptionally elevated or decreased risk of Atrial Fibrillation, we create several updates to the Exceptional Model Mining (EMM) [16, 35] framework (cf. Section 2.2) that allow us to analyze the ECG signals as targets. The patient characteristics are the *descriptors*, the extracted ECG features are the *targets*, and AF complications serve as *evaluators*. The *model class* is ECG morphology abnormalities in the form of phenotypes.

5.1 Quality Measures

We derive multiple quality measures to assess multiple kinds of exceptional behavior in ECGs.

5.1.1 Heart Rate Variability, P-wave Absence, F-wave Presence. First, we define some statistical functions θ based on ECG phenotypes related to the Heart Rate Variability (HRV) and P-waves extracted in Section 4.1. We implement three measures (identified in AF detection methods mentioned in Section 2) to find exceptionality in the HRV: Standard Deviation of all RR-intervals (SDRR), Root Mean Square of Successive RR-interval Differences (RMSSD) and Standard Deviation of Successive RR-interval Differences (SDSD).

$$\theta_{\text{SDRR}}(p) = \sqrt{\frac{1}{K-2}\sum_{i=1}^{K-1}(RR_i - \overline{RR})^2}$$

$$\theta_{\text{RMSSD}}(p) = \sqrt{\frac{1}{K-3}\sum_{i=1}^{K-2}(\Delta RR_i)^2}$$

$$\theta_{\text{SDSD}}(p) = \sqrt{\frac{1}{K-3}\sum_{i=1}^{K-2}(\Delta RR_i - \overline{\Delta RR})^2}$$

Here, K is the number of heartbeats, and $\overline{RR}$ and $\overline{\Delta RR}$ are the average RR-interval and average difference between consecutive RR-intervals, respectively; p denotes the patient. We also record the percentage of heartbeats without P-waves, denoted $\theta_P(p)$, and the percentage of heartbeats displaying F-waves, denoted $\theta_F(p)$.

5.1.2 Combining HRV with P-waves or F-waves. The arrhythmias in Table 2 can be distinguished by a combination of 1) observations on HRV, and 2) P-wave absence or F-wave presence. The interest lies in finding cases where the patient experiences high HRV while either no P-wave is present in the heartbeat or replacement F-waves occur. Due to the sequential steps of our feature extraction procedure, directly matching the RR-intervals with the P-waves and/or F-waves of the same heartbeat is a bad idea. RR-intervals induce a list of S- and Q-peaks. If unmatched, S- and/or Q-peaks are filtered until only proper SQ-intervals remain. These then in turn induce the discovery of P-waves and F-waves. This makes it trivial to match P-waves and F-waves to the correct SQ-interval duration; the link with the correct RR-interval duration may be less obvious. Hence, combining P- and/or F-wave observations with RR-intervals may mislead, and hence, the following choices for θ are computed with the SQ-interval measures in place of the RR-interval.

We adapt SDRR, RMSSD, and SDSD such that they seek the combination of exceptional SQ-intervals (instead of RR) and either the absence of P-waves or the presence of F-waves. This leads to six additional choices for θ.

$$\theta^*_{\text{SDSQ}}(p) = \sqrt{\frac{1}{K-2}\sum_{i=1}^{K-1}\mathbb{1}_*(p,i)\cdot(SQ_i - \overline{SQ})^2}$$

$$\theta^*_{\text{RMSSD}}(p) = \sqrt{\frac{1}{K-3}\sum_{i=1}^{K-2}\mathbb{1}_*(p,i)\cdot(\Delta SQ_i)^2}$$

$$\theta^*_{\text{SDSD}}(p) = \sqrt{\frac{1}{K-3}\sum_{i=1}^{K-2}\mathbb{1}_*(p,i)\cdot(\Delta SQ_i - \overline{\Delta SQ})^2}$$

Here, $* \in \{\text{P-wave absent, F-wave present}\}$, and $\mathbb{1}$ is the indicator function (equals 1 if its subscript clause is true for heartbeat i of patient p, equals 0 if false). Recall that both the absence of P-waves and the presence of F-waves are indicators of AF.

5.1.3 Correcting for Subgroup Size and Precision. If we let the subgroup search be guided by any of the θs defined in the previous section, the search strategy will run towards tiny subgroups covering an insubstantially small part of the data set. This is a common problem in local pattern mining: without correction for subgroup size, tiny subgroups will dominate the result set. Several solutions to this problem exist; we incorporate the entropy (denoted $\varphi_{\text{ef}}(D)$, cf. [16, Section 3.2.1]) of the split between subgroup and complement as a factor in our quality measure. Having made this correction, the search strategy discovers substantially larger subgroups, but these now also include more negatives (patients who didn't experience AF). This is not surprising: since we look at recorded signals before AF, preceding phenotypes are more subtle. To enable finding weaker phenotypes, we incorporate precision (denoted $\varphi_{\text{pr}}(D)$, cf. [19, Section 4.3]) as a factor in our quality measure. This encourages True Positives (TP) over False Positives (FP).

5.1.4 Bringing it All Together. For any of the eleven θs from Sections 5.1.1 and 5.1.2, we can compute its average $\bar{\theta}^X$ for any subset $X \subseteq \Omega$. We derive the phenotype-related $\varphi_{\text{pheno}}(D) = \bar{\theta}^{G_D} - \bar{\theta}^{\Omega}$. Finally, we create a compound QM that rewards a subgroup if it scores high on all of the phenotype, entropy, and precision components.

$$\varphi(D) = \varphi_{\text{ef}}(D) \cdot \varphi_{\text{pr}}(D) \cdot \varphi_{\text{pheno}}(D)$$

5.2 Generating Candidate Subgroups

We employ an existing search strategy for supervised LPM: *beam search* [16, Algorithm 1]. Candidate subgroups are generated in a level-wise, general-to-specific manner. On the first level, we loop over the descriptors, generating all sensible single conditions on single attributes. Evaluating them with φ, we keep a preset number w (the *beam width*) of the best-scoring candidates as the *level-1 beam*. We generate candidates for each subsequent level i by looping over the subgroups in the level-$(i-1)$ beam, and *refining*[1] them by conjoining additional conditions, generated by looping over the attributes; after candidate generation, the best w are kept as the level-i beam. Beam search terminates after a predefined maximum level d (the *search depth*); the top-q subgroups encountered along the way are reported.

5.3 Anti-Redundancy Methods

Any EMM search strategy runs the risk of returning a top-q of very similar, exceptional subgroups: near copies of very exceptional subgroups are likely also very exceptional. This may drown out other subgroups that are also exceptional. To tackle this problem, the EMM literature contains several anti-redundancy methods.

We remove *similarity redundancy*: identical copies of descriptions whose conjuncts appear in a different order. We also remove *generalization redundancy* through description-based selection [34, 53]: if two overlapping descriptions, e.g. "age $\leq$ 50" and "age $\leq$ 50 AND smokes = no", have the same quality score, the more general description is kept.

[1] using the refinement operator η from [16, Section 4.1], parameterized with the `1bca` numeric refinement strategy from [45], and augmented with the ordinal refinement operator from [53].

Table 3: The number and percentage of patients in Ω that experienced the various types of AF (groups overlap; unfortunate patients may experience AF in multiple stages).

	AF Complication identified	
Stages	**Yes**	**No**
before	6 (2.6%)	224
during	6 (2.6%)	224
after	75 (32.6%)	155
future	1 (0.4%)	229
AF in general	87 (37.8%)	143

6 Experiments

Our main experiments involve proprietary hospital data. For reproducibility purposes, we repeat the experiment on a public data set; details and results can be found in the Supplementary Material at [9, Section 2]. Source code and more material are publicly available on GitHub[2]; our implementation builds upon the pseudocode from [16, Algorithm 1] and the implementation from [53].

6.1 Data

The data set consists of 230 patients who have undergone cardiac surgery. In total, 186 men (80.9%) and 44 women (19.1%) are included, risking bias. However, men are more prone to heart failure, thus needing cardiac surgery, than women [17]: it is a natural consequence unlikely to affect our research. Our data set includes Atrial Fibrillation (AF) complications experienced before, during, and/or after an operation, summarized in Table 3. Postoperative AF occurs in the first four weeks after surgery [44]; any AF detected later than these four weeks is considered an unrelated consequence.

The data set includes lead II ECG recordings, Electronic Health Records (EHR), and AF complication indicators. Due to the nature of the data, each entry is assumed to be a new patient; for all practical purposes, this is close enough to the truth such that the probability becomes negligible that violations of this assumption substantially affect the results of our analysis.

After transforming the EHR with aggregation and one-hot encoding, we end up with 247 medical characteristics as descriptors. Subsequent filtering based on domain expertise is undesired due to the exploratory nature of beam search.

6.2 Experimental Setup

The beam search parameters are carefully chosen to combine a wide exploration of the descriptor space while keeping results clinically relevant. The search depth $d = 3$ ensures transparency and comprehensibility of the resulting descriptions; exploring conjunctions of more descriptors would be easily computationally feasible, but the resulting subgroups would be harder to interpret and more likely to be false discoveries. The beam width $w = 50$ allows for over 25% of the descriptors to participate in the level-1 beam, and keeps many potentially interesting descriptions as long as possible; we sacrifice runtime for a wider exploration of the descriptor space. The top-q descriptions to be returned is set to $q = 15$. Finally, to

[2] https://github.com/liekevandenbiggelaar/EAFM.

stimulate generality in the evidence of exceptional behavior, we set the minimum coverage of any subgroup to 5% of $|\Omega|$.

We judge the validity and medical relevance of our discovered subgroups by evaluating the content with domain experts. We round all quality scores φ to two decimals. Under this choice, we have observed identical qualities only for subgroups covering the exact same set of patients, where it is to be expected.

6.3 Results

Table 4 displays summary statistics of the results. A good subgroup has high exceptionality of the phenotype and a high percentage of AF. All experiments find more patients with AF than without, and a high exceptionality in phenotypes. Experiments on individual phenotypes (# 1–5) yield an average AF rate above 70% (while the AF rate on Ω is 37.8%). However, they (except for θ_F) score lower on the phenotype exceptionality than most combined phenotypes (# 6–11, except for θ^P_{SDSQ} and θ^F_{SDSD}): those find subgroups whose phenotype is over 2.5 times more exceptional than the population. The commonality of variations in phenotypes could cause this: as many patients have one of the variational properties, the precision factor φ_{pr} has a greater effect. Hence, Experiments 1-5 find descriptions of patients who often have AF, but not necessarily exceptional ECG phenotypes, as was the objective. Experiments 6-11 find subgroups encompassing both ECG phenotypes differing from Ω and the majority of AF patients. Some episodes of AF are likely missed due to the way ECG monitoring is currently handled; improvements in this monitoring and postprocessing of the ECG signal are promising avenues of future work in order to also cover the remaining AF patients.

The remainder of this section will focus on the experiments with combined phenotypes only. Although Experiment 5 on F-waves seems interesting from the discussion in the previous paragraph, extracting these waves from the ECG is more an art than an exact science; manual choices will influence results. The use of only this phenotype will likely lead to biased results focusing on outlier cases. Therefore, we include F-waves only in combination with high heart rate variability.

The full top-15 subgroups reported by beam search for each of the eleven experiments are listed in the Supplementary Material [9, Section 1]. We discussed and postprocessed the medically most relevant findings with medical doctors at the hospital; the remainder of this section presents the results of that discussion, involving experiments with θ^P_{SDSD}, θ^P_{RMMSD}, and θ^F_{SDSD}.

6.3.1 Subgroups Discovered with θ^P_{SDSD}. The θ^P_{SDSD} phenotype seeks patients with varying heart rate related to the standard deviation of successive SQ-interval differences, and a high percentage of missing P-waves. The top chart in Figure 2 shows that four subgroups have a low percentage of AF and five subgroups have low phenotype exceptionality. Leaving out one clinically uninteresting subgroup, Table 5 lists the five remaining subgroups.

All subgroups select patients with acidic blood (normal/high *standard bicarbonate*, low *anion gap*, and high *chloride*). Three select patients with blood clotting problems (high *prothrombin time*), who are likely to be assisted by the heart-lung machine. A combination of these two occurs three times. In two other situations, acidic blood is combined with patients with prediabetes (normal/high *glucose*).

Table 4: Summary statistics of the eleven experiments: experiment number, phenotype, average φ_{pheno} of the top-15 subgroups, average phenotype over the full data set, phenotype exceptionality factor, and percentage of patients in the subgroups that display AF (for the full data set: 37.8%**).**

#	Phenotype	$\bar{\varphi}_{\text{pheno}}$	$\bar{\theta}^{\Omega}$	$\frac{\bar{\varphi}_{\text{pheno}}}{\bar{\theta}^{\Omega}}$	% AF
1	θ_{SDSD}	103.38	75.38	1.37	**0.71**
2	θ_{RMSSD}	141.38	93.92	1.51	**0.71**
3	θ_{SDRR}	113.38	81.63	1.39	**0.73**
4	θ_P	50.48	30.38	1.66	**0.71**
5	θ_F	1.88	0.54	**3.48**	**0.70**
6	θ^P_{SDSD}	490.56	183.52	**2.67**	0.60
7	θ^P_{RMSSD}	752.20	277.89	**2.71**	0.58
8	θ^P_{SDSQ}	656.41	396.64	1.65	0.57
9	θ^F_{SDSD}	410.66	134.76	**3.05**	0.59
10	θ^F_{RMSSD}	752.20	277.89	**2.71**	0.58
11	θ^F_{SDSQ}	657.03	426.43	1.54	0.58

The risk groups that emerge from this experiment are patients with: (A) acidic blood and blood clotting problems (#1, #9, #11); (B) acidic blood and assistance by the heart-lung machine (#1, #9, #11); (C) acidic blood and prediabetes (#5, #6). These results are partially supported by existing literature. AF risk is elevated in patients with a low anion gap [22], and prediabetes [27]. Furthermore, low chloride levels increase the risk of AF, but high chloride levels, as in our subgroups, show a risk similar to regular levels [21].

6.3.2 Subgroups Discovered with θ^P_{RMSSD}. The θ^P_{RMSSD} phenotype seeks patients with varying heart rate related to the root mean square of successive RR-interval differences and a high percentage of missing P-waves. The middle chart in Figure 2 shows that five subgroups have a low percentage of AF and four subgroups have low phenotype exceptionality. Four other subgroups were already discovered with θ^P_{SDSD}. Table 5 lists the two remaining subgroups.

The one new subgroup strengthens our hypothesis from the θ^P_{SDSD} experiment, on patients combining acidic blood (high *chloride*) with either blood clotting problems (high *prothrombin time*) or heart-lung machine assistance. The other new subgroup suggests exceptionality in patients receiving antibiotics (*cefazolin* treatment).

The emerging risk groups are patients with: (A) blood clotting problems and antibiotics administration in the form of cefazolin (#5); (B) acidic blood and blood clotting problems (#13); (C) acidic blood and assistance by the heart-lung machine (#13). The administration of antibiotics has been associated with an increased risk of AF [3], but cefazolin is not yet mentioned in the list of antibiotics tested. As in the θ^P_{SDSD} experiment, high chloride is named, which contraindicates results from existing literature [21].

6.3.3 Subgroups Discovered with θ^F_{SDSD}. The θ^F_{SDSD} phenotype seeks patients with varying heart rate related to standard deviation of successive SQ-interval differences and a high percentage of replacement F-waves. The bottom chart in Figure 2 shows that four subgroups have a low AF percentage and four have low phenotype

Table 5: Subgroups discovered when seeking exceptional phenotypes θ^P_{SDSD}, θ^P_{RMSSD}, and θ^F_{SDSD}. Raw values are postprocessed into qualitative statements, guided by medical professionals and medical literature.

Phenotype	#	Description
θ^P_{SDSD}	1	high *prothrombin time* ∧ normal/high *standard bicarbonate* ∧ no *thrombin* treatment
θ^P_{SDSD}	5	low *anion gap* ∧ normal/high *glucose* ∧ no *minimally invasive aortic valve replacement*
θ^P_{SDSD}	6	low *anion gap* ∧ normal/high *glucose* ∧ *sex* is male
θ^P_{SDSD}	9	high *prothrombin time* ∧ normal/high *standard bicarbonate* ∧ high *chloride*
θ^P_{SDSD}	11	high *prothrombin time* ∧ normal/high *standard bicarbonate* ∧ normal *sodium*
θ^P_{RMSSD}	5	high *prothrombin time* ∧ *cefazolin* treatment ∧ normal *FiO2*
θ^P_{RMSSD}	13	high *prothrombin time* ∧ little *blood loss* ∧ high *chloride*
θ^F_{SDSD}	11	high *prothrombin time* ∧ *Ringer's lactate* treatment ∧ normal *FiO2*
θ^F_{SDSD}	14	high *prothrombin time* ∧ normal/high *standard bicarbonate* ∧ normal *potassium*

exceptionality. Five other subgroups were already discovered with θ^P_{SDSD} or θ^P_{RMSSD}. Table 5 lists the two remaining subgroups.

Again, the overlapping subgroups complement our previous findings. One subgroup also includes patients with acidic blood (normal/high *standard bicarbonate*) and blood clotting problems (high *prothrombin time*). The remaining subgroup includes patients with blood clotting problems and a low blood volume/blood pressure (*Ringer's lactate* treatment).

The emerging risk groups are patients with: (A) blood clotting problems and low blood volume/blood pressure (#11); (B) acidic blood and blood clotting problems (#14); (C) acidic blood and assistance by the heart-lung machine (#14). Although previous research has found a higher prevalence of AF after the administration of Ringer's lactate [52], this treatment had not yet been associated with increased risk.

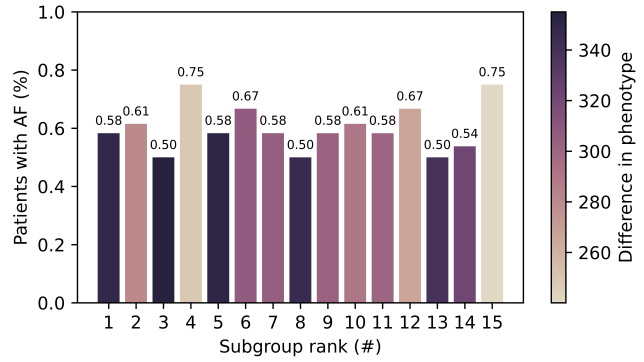
7 Conclusions

Cardiac surgery puts patients at risk of Atrial Fibrillation (AF) [44]. This heart arrhythmia is the leading cardiac cause of strokes [18]. The consequences can be avoided by preventive treatment, if we can see coming that AF is likely to occur. Current practice is reactive: medical professionals can only act when the episode is already occurring, signaled by an alarm based on unknown factors. Instead, we provide an instance of the Exceptional Model Mining (EMM) [16, 35] framework, discovering transparent and actionable subgroups of patients at a higher risk of AF. We do so by processing ECG data from patients into AF-related phenotypes, and defining quality measures that reward subgroups combining exceptional phenotype behavior, a high percentage of AF occurrence, and a substantial size. The subgroups found are hypothesized to have a higher risk of developing AF after surgery, and our advice to the hospital is to give preventive medications to patients who match the descriptions found. As we work with sensitive data, ethical considerations should be taken into account. Therefore, our research should be regarded as exploratory rather than confirming. Ideally, the knowledge derived from this paper should be confirmed in follow-up medical studies. In that sense, the contributions of this paper are a step towards stratified medicine, where EMM automatically discovers hypotheses for interesting strata to explore further.

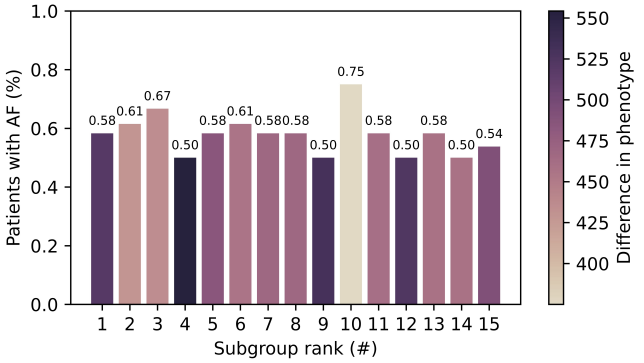
The method is in deployment at the Catharina Hospital in Eindhoven, the Netherlands. We conduct several experiments to select the best method for the deployment. All methods discover exceptional subgroups in terms of phenotypes and AF. Methods relying solely on heart rate variability (HRV) measurements, P-wave absence, or F-wave presence deliver subgroups with the strongest AF incidence rate. Still, the phenotypes are not that different from the overall patient population in our data set. Methods relying on combinations of on the one hand HRV and on the other hand either P-wave absence or F-wave presence, have a slightly less strong but still substantially elevated AF incidence rate, combined with strongly deviating phenotypes (cf. Table 4). Some of the subgroups we discover are defined by characteristics associated with an increased risk of AF in existing literature, thereby confirming the validity of our findings. Other subgroups we discover represent new hypotheses in this area for AF risk groups. The advice derived from the model is to administer preventive medication to prevent AF from occurring in the risk groups found and execute follow-up medical studies on the relationship of the following characteristics with the risk of AF: 1) patients assisted by the heart-lung machine that also have acidic blood; 2) patients with high chloride levels; 3) patients with blood clotting problems that need cefazolin admission to prevent infection; 4) patients with blood clotting problems that need Ringer's lactate to overcome low blood volume/pressure.

7.1 Future Work

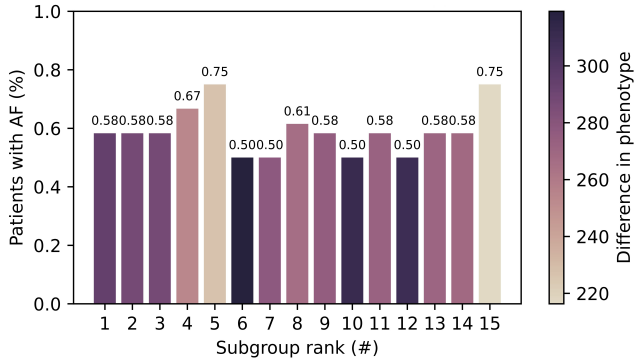
Our experiments run on data from one of the twelve leads for measuring ECG signals. The cardiac signal preprocessing of Section 4 relies on this choice: including other leads harms the current procedure (interference, more noise), which would require more complex preprocessing steps. Since intelligently combining data from multiple leads holds the potential to provide more subtle information on the heart health of the patient, including more leads is an avenue to be explored in future work.



(a) Experiment #6: θ^{P}_{SDSD}.



(b) Experiment #7: $\theta^{P}_{\text{RMSSD}}$.



(c) Experiment #9: θ^{F}_{SDSD}.

Figure 2: A visual representation of the subgroups found with the three selected experiments. The bar height shows the percentage of patients with AF, while color indicates phenotype differences from the population average.

Our method is in deployment on the data from a single hospital, the Catharina Hospital in Eindhoven, the Netherlands. The method ought to be deployable more generally, on data from other hospitals in other locations. The only requirement for such deployment to take place, is that the hospital must possess data that comes in the shape as illustrated in Table 1: all three modalities of this data (EHR descriptors, ECG targets, AF complication occurrence annotations) must be present. We are open to collaborations with hospitals where such data (with all three modalities) exists. However, we made our code publicly available deliberately: if such a hospital would decide to deploy our method without our direct involvement, this would of course also be a most welcome development.

Finally, the deployment of EMM on the data from the Catharina Hospital resulted in the discovery of subgroups of patients at an elevated AF risk. Some of these were known risk groups, others form new hypotheses with the potential for better medical intervention in the future. We list four specific subgroups that represent candidates for medical follow-up studies in the paragraph preceding Section 7.1. Meanwhile, we are setting up a clinical follow-up study on whether the administration of Alfentanil (which occurs in multiple discovered subgroups across multiple AF phenotypes; cf. [9, Tables 1–11]) can be replaced by other opiates including but not limited to Remifentanil.

References

[1] A.M.E. Abdelgawad, M.A. Hussein, H. Naeim, R. Abuelatta, and S. Alghamdy. 2019. A comparative study of TAVR versus SAVR in moderate and high-risk surgical patients: Hospital outcome and midterm results. *The Heart Surgery Forum* 22, 5 (2019), E331—E339. doi:10.1532/hsf.2243

[2] S. Abhishek, S. Veni, and K.A. Narayanankutty. 2019. Biorthogonal wavelet filters for compressed sensing ECG reconstruction. *Biomedical Signal Processing and Control* 47 (2019), 183–195. doi:10.1016/j.bspc.2018.08.011

[3] E. Abo-Salem, J.C. Fowler, M. Attari, C.D. Cox, A. Perez-Verdia, R. Panikkath, and K. Nugent. 2014. Antibiotic-induced cardiac arrhythmias. *Cardiovascular Therapeutics* 32, 1 (2014), 19–25. doi:10.1111/1755-5922.12054.

[4] R. Agrawal and S. Ramakrishnan. 1994. Fast algorithms for mining association rules in large databases. In *VLDB '94: Proceedings of the 20th International Conference on Very Large Data Bases*. VLDB, 487–499. doi:10.5555/645920.672836

[5] Z.I. Attia, P.A. Noseworthy, F. Lopez-Jimenez, S.J. Asirvatham, A.J. Deshmukh, B.J. Gersh, R.E. Carter, X. Yao, A.A. Rabinstein, B.J. Erickson, et al. 2019. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: A retrospective analysis of outcome prediction. *The Lancet* 394, 10201 (2019), 861–867. doi:10.1016/S0140-6736(19)31721-0

[6] M. Atzmueller. 2015. Subgroup discovery. *WIREs Data Mining Knowl. Discov.* 5, 1 (2015), 35–49. doi:10.1002/WIDM.1144

[7] G. Bényász and L. Cser. 2013. Wavelet based de-noising in manufacturing and in business. *Procedia CIRP* 12 (2013), 282–287. doi:10.1016/j.procir.2013.09.049

[8] G. Bianchi and R. Sorrentino. 2007. *Electronic filter simulation & design.* McGraw-Hill Education.

[9] L. van den Biggelaar, R.M. Schouten, A. de Bie, R.A. Bouwman, and W. Duivesteijn. 2025. Supplementary Material to "Characterizing the Risk of Atrial Fibrillation in Cardiac Patients with Exceptional Electrocardiogram Phenotypes". (2025). doi:10.6084/m9.figshare.29135480

[10] K.H. Boon, M. Khalil-Hani, and M.B. Malarvili. 2018. Paroxysmal atrial fibrillation prediction based on heart rate variability analysis and non-dominated sorting genetic algorithm III. *Computer Methods and Programs in Biomedicine* 153 (2018), 171–184. doi:10.1016/j.cmpb.2017.10.012

[11] W. Cai, Y. Chen, J. Guo, B. Han, Y. Shi, L. Ji, J. Wang, G. Zhang, and J. Luo. 2020. Accurate detection of atrial fibrillation from 12-lead ECG using deep neural network. *Computers in Biology and Medicine* 116 (2020), 103378. doi:10.1016/j.compbiomed.2019.103378

[12] C.C. Chen and F.R. Tsui. 2020. Comparing different wavelet transforms on removing electrocardiogram baseline wanders and special trends. *BMC Medical Informatics and Decision Making* 20, 343 (2020). doi:10.1186/s12911-020-01349-x

[13] J. Christie. 2008. One size does not fit all. *Gastrointestinal Endoscopy* 67, 2 (2008), 278–279. doi:10.1016/j.gie.2007.10.007.

[14] L. Clavier, J.M. Boucher, R. Lepage, J.J. Blanc, and J.C. Cornily. 2002. Automatic P-wave analysis of patients prone to atrial fibrillation. *Medical and Biological Engineering and Computing* 40, 1 (2002), 63–71. doi:10.1007/BF02347697

[15] M.B. Conover. 2003. *Understanding electrocardiography.* Elsevier Health Sciences, Chapter 1. The 12 electrocardiogram leads, 3–22.

[16] W. Duivesteijn, A. Feelders, and A. Knobbe. 2016. Exceptional model mining: Supervised descriptive local pattern mining with complex target concepts. *Data Mining and Knowledge Discovery* 30 (2016), 47–98. doi:10.1007/s10618-015-0403-4

[17] A. Ehdaie, E. Cingolani, M. Shehata, X. Wang, A.B. Curtis, and S.S. Chugh. 2018. Sex differences in cardiac arrhythmias: Clinical and research implications. *Circulation: Arrhythmia and Electrophysiology* 11, 3 (2018), e005680. doi:10.1161/CIRCEP.117.005680

[18] K.L. Furie, L.B. Goldstein, G.W. Albers, P. Khatri, R. Neyens, M.P. Turakhia, T.N. Turan, and K.A. Wood. 2012. Oral antithrombotic agents for the prevention of stroke in nonvalvular atrial fibrillation: A science advisory for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 43, 12 (2012), 3442–3453. doi:10.1161/STR.0b013e318266722a

[19] J. Fürnkranz and P.A. Flach. 2005. Roc 'n'rule learning—towards a better understanding of covering algorithms. *Machine Learning* 58 (2005), 39–77. doi:0.1007/s10994-005-5011-x

[20] D. Gamberger, N. Lavrač, A. Krstačić, and G. Krstačić. 2007. Clinical data analysis based on iterative subgroup discovery: Experiments in brain ischaemia data analysis. *Applied Intelligence* 27, 3 (2007), 205–217. doi:10.1007/s10489-007-0068-9

[21] Y. Gan, S. Nie, M. Pang, R. Huang, H. Xu, B. Liu, J. Weng, C. Chunbo, H. Liu, H. Li, et al. 2024. Inverse association between serum chloride levels and the risk of atrial fibrillation in chronic kidney disease patients. *Clinical Kidney Journal* 17, 8 (2024), sfae137. doi:10.1093/ckj/sfae137

[22] S. Geurts, M.J. Tilly, B. Arshi, B.H.C. Stricker, J.A. Kors, J.W. Deckers, N.M.S. de Groot, M.A. Ikram, and M. Kavousi. 2023. Heart rate variability and atrial fibrillation in the general population: A longitudinal and Mendelian randomization study. *Clinical Research in Cardiology* 112, 6 (2023), 747–758. doi:10.1007/s00392-022-02072-5.

[23] S. Guo, X. Li, H. Liu, P. Zhang, X. Du, G. Xie, and F. Wang. 2017. Integrating temporal pattern mining in ischemic stroke prediction and treatment pathway discovery for atrial fibrillation. In *AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science*, Vol. 2017. AMIA, 122–130. https://pmc.ncbi.nlm.nih.gov/articles/PMC5543383/

[24] H. Henderi, T. Wahyuningsih, and E. Rahwanto. 2021. Comparison of min-max normalization and Z-score normalization in the k-nearest neighbor algorithm to test the accuracy of types of breast cancer. *International Journal of Informatics and Information Systems* 4, 1 (2021), 13–20. doi:10.47738/ijiis.v4i1.73

[25] F. Herrera, C.J. Carmona, P. González, and M.J. del Jesus. 2011. An overview on subgroup discovery: foundations and applications. *Knowledge and Information Systems* 29, 3 (2011), 495–525. doi:10.1007/S10115-010-0356-2

[26] K. Ishida, H. Hayashi, A. Miyamoto, Y. Sugimoto, M. Ito, Y. Murakami, and M. Horie. 2010. P wave and the development of atrial fibrillation. *Heart Rhythm* 7, 3 (2010), 289–294. doi:10.1016/j.hrthm.2009.11.012

[27] C. Johansson, L. Örtendahl, M. M Lind, J. Andersson, L. Johansson, and M. Brunström. 2023. Diabetes, prediabetes, and atrial fibrillation: A population-based cohort study based on national and regional registers. *Journal of Internal Medicine* 294, 5 (2023), 605–615. doi:10.1111/joim.13688

[28] R.M. John and S. Kumar. 2016. Sinus node and atrial arrhythmias. *Circulation* 133, 19 (2016), 1892–1900. doi:10.1161/CIRCULATIONAHA.116.018011

[29] S.J. Jung, C.S. Son, M.S. Kim, D.J. Kim, H.S. Park, and Y.N. Kim. 2013. Association rules to identify complications of cerebral infarction in patients with atrial fibrillation. *Healthcare Informatics Research* 19, 1 (2013), 25–32. doi:10.4258/hir.2013.19.1.25

[30] R. Kher. 2019. Signal processing techniques for removing noise from ECG signals. *Journal for Biomedical Engineering* 3, 101 (2019), 1–9. doi:10.17303/jber.2019.3.101

[31] W. Klösgen. 1996. Explora: A multipattern and multistrategy discovery assistant. In *Advances in Knowledge Discovery and Data Mining*. AAAI, 249–271. doi:10.5555/257938.257965

[32] J. Kornej, C.S. Börschel, E.J. Benjamin, and R.B. Schnabel. 2020. Epidemiology of atrial fibrillation in the 21st century: Novel methods and new insights. *Circulation Research* 127, 1 (2020), 4–20. doi:10.1161/CIRCRESAHA.120.316340

[33] J. Lee, Y. Nam, D.D. McManus, and K.H. Chon. 2013. Time-varying coherence function for atrial fibrillation detection. *IEEE Transactions on Biomedical Engineering* 60, 10 (2013), 2783–2793. doi:10.1109/TBME.2013.2264721

[34] M. van Leeuwen and A. Knobbe. 2012. Diverse subgroup set discovery. *Data Mining and Knowledge Discovery* 25 (2012), 208–242. doi:10.1007/s10618-012-0273-y

[35] D. Leman, A. Feelders, and A. Knobbe. 2008. Exceptional model mining. In *Machine Learning and Knowledge Discovery in Databases*, Walter Daelemans, Bart Goethals, and Katharina Morik (Eds.). Springer, Berlin, Heidelberg, 1–16. doi:10.1007/978-3-540-87481-2_1

[36] H. Li and P. Boulanger. 2021. An automatic method to reduce baseline wander and motion artifacts on ambulatory electrocardiogram signals. *Sensors* 21, 24 (2021), 8169. doi:10.3390/s21248169

[37] J. Li, M. Gao, M. Zhang, D. Liu, Z. Li, J. Du, and Y. Hou. 2020. Treatment of atrial fibrillation: A comprehensive review and practice guide. *Cardiovascular Journal of Africa* 31, 3 (2020), 153–158. doi:10.5830/CVJA-2019-064

[38] Q. Li, Y. Zhang, H. Kang, Y. Xin, and C. Shi. 2017. Mining association rules between stroke risk factors based on the Apriori algorithm. *Technology and Health Care: Official Journal of the European Society for Engineering and Medicine* 25, S1 (2017), 197–205. doi:10.3233/THC-171322

[39] X. Li, H. Liu, X. Du, G. Hu, G. Xie, and P. Zhang. 2016. Using frequent item set mining and feature selection methods to identify interacted risk factors: The atrial fibrillation case study. In *Studies in Health Technology and Informatics*. Vol. 228. IOS Press, 562–566. doi:10.3233/978-1-61499-678-1-562

[40] L.A. Lopes and D.K. Agrawal. 2022. Post-operative atrial fibrillation: Current treatments and etiologies for a persistent surgical complication. *Journal of Surgery and Research* 5, 1 (2022), 159–172. doi:0.26502/jsr.10020209

[41] E.J.S. Luz, W.B. Schwartz, G. Cámara-Chávez, and D. Menotti. 2016. ECG-based heartbeat classification for arrhythmia detection: A survey. *Computer Methods and Programs in Biomedicine* 127 (2016), 144–164. doi:10.1016/j.cmpb.2015.12.008

[42] D. Makowski, T. Pham, Z.J. Lau, J.C. Brammer, F. Lespinasse, H. Pham, C. Schölzel, and A.S.H. Chen. 2021. NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behavior Research Methods* 53 (2021), 1689–1696. doi:10.3758/s13428-020-01516-y

[43] J.P. Martinez, R. Almeida, S. Olmos, A.P. Rocha, and P. Laguna. 2004. A wavelet-based ECG delineator: Evaluation on standard databases. *IEEE Transactions on Biomedical Engineering* 51, 4 (2004), 570–581. doi:10.1109/TBME.2003.821031

[44] W.F. McIntyre. 2023. Post-operative atrial fibrillation after cardiac surgery: Challenges throughout the patient journey. *Frontiers in Cardiovascular Medicine* 10 (2023), 1156626. doi:10.3389/fcvm.2023.1156626

[45] M. Meeng and A.J. Knobbe. 2021. For real: A thorough look at numeric attributes in subgroup discovery. *Data Mining and Knowledge Discovery* 35, 1 (2021), 158–212. doi:10.1007/s10618-020-00703-x

[46] K. Morik, J.F. Boulicaut, and A. Siebes. 2005. *Local pattern detection*. Springer, Berlin, Heidelberg. doi:10.1007/b137601

[47] H. Ogawa, Y. An, H. Nishi, S. Fukuda, K. Ishigami, S. Ikeda, K. Doi, Y. Ide, Y. Hamatani, A. Fujino, et al. 2021. Characteristics and clinical outcomes in atrial fibrillation patients classified using cluster analysis: The Fushimi atrial fibrillation registry. *EP Europace* 23, 9 (2021), 1369–1379. doi:10.1093/europace/euab079

[48] J. Oster and G.D. Clifford. 2015. Impact of the presence of noise on RR interval-based atrial fibrillation detection. *Journal of Electrocardiology* 48, 6 (2015), 947–951. doi:10.1016/j.jelectrocard.2015.08.013

[49] V. Pandey and V.K. Giri. 2016. High frequency noise removal from ECG using moving average filters. In *2016 International Conference on Emerging Trends in Electrical Electronics & Sustainable Energy Systems (ICETEESES)*. IEEE, 191–195. doi:10.1109/ICETEESES.2016.7581383

[50] A. Parsi, M. Glavin, E. Jones, and D. Byrne. 2021. Prediction of paroxysmal atrial fibrillation using new heart rate variability features. *Computers in Biology and Medicine* 133 (2021), 104367. doi:10.1016/j.compbiomed.2021.104367

[51] S. Petrutiu, J. Ng, G.M. Nijm, H. Al-Angari, S. Swiryn, and A.V. Sahakian. 2006. Atrial fibrillation and waveform characterization: A time domain perspective in the surface ECG. *IEEE engineering in medicine and biology magazine : the quarterly magazine of the Engineering in Medicine & Biology Society* 25, 6 (2006), 24–30. doi:10.1109/emb-m.2006.250505

[52] S. Redant, Y. Langman, D. De Bels, R. Attou, and P.M. Honore. 2019. Mechanism of arrhythmias during the infusion of Ringer's acetate and Ringer's lactate solutions during cardiac surgery: New insights. *Critical Care* 23, 1 (2019), 413. doi:10.1186/s13054-019-2696-y

[53] R.M. Schouten, W. Duivesteijn, and M. Pechenizkiy. 2022. Exceptional model mining for repeated cross-sectional data (EMM-RCS). In *Proceedings of the 2022 SIAM International Conference on Data Mining (SDM)*. SIAM, 585–593. doi:10.1137/1.9781611977172.66

[54] K. Tateno and L. Glass. 2000. A method for detection of atrial fibrillation using RR intervals. In *Computers in Cardiology*, Vol. 27. IEEE, 391–394. doi:10.1109/CIC.2000.898539

[55] M.R. Trusheim, E.R. Berndt, and F.L. Douglas. 2007. Stratified medicine: Strategic and economic implications of combining drugs and clinical biomarkers. *Nature Reviews Drug Discovery* 6, 4 (2007), 287–293. doi:10.1038/nrd2251

[56] M. Vitolo, M. Proietti, A. Shantsila, G. Boriani, and G.Y.H. Lip. 2021. Clinical phenotype classification of atrial fibrillation patients using cluster analysis and associations with trial-adjudicated outcomes. *Biomedicines* 9, 7 (2021), 843. doi:10.3390/biomedicines9070843

[57] S. Wrobel. 1997. An algorithm for multi-relational discovery of subgroups. In *Principles of Data Mining and Knowledge Discovery, First European Symposium, PKDD '97, Trondheim, Norway, June 24-27, 1997, Proceedings (Lecture Notes in Computer Science, Vol. 1263)*. Springer, Berlin, Heidelberg, 78–87. doi:10.1007/3-540-63223-9_108

[58] Y. Xia, N. Wulan, K. Wang, and H. Zhang. 2018. Detecting atrial fibrillation by deep convolutional neural networks. *Computers in Biology and Medicine* 93 (2018), 84–92. doi:10.1016/j.compbiomed.2017.12.007