Improving Long QT Syndrome diagnosis using machine learning on ECG characteristics

Job Stoks
MSc Thesis
Technical Medicine
Faculty of Science and Technology
University of Twente

April 4, 2018

UNIVERSITY
OF TWENTE.

university hospital
Maastricht
aZM

UM Maastricht University

Maastricht UMC⁺

Abstract

Introduction Congenital long QT syndrome (LQTS) is a genetic disorder affecting cardiac ion channels which leads to an increased risk of malignant ventricular arrhythmias and sudden cardiac death [1]. Diagnosing LQTS remains challenging because of a considerable overlap of the QT-interval between LQTS patients and healthy controls [2]. Analysis of T-wave morphology has shown to be of discriminative value to diagnose LQTS [3–6]. An objective diagnostic tool that includes T-wave morphology might further improve LQTS diagnosis.

Methods and results A retrospective study was performed on 699 standard ECGs recorded from patients with LQT1, LQT2 and LQT3 and genotype-negative relatives. T-wave morphology parameters and subject characteristics were used as inputs to three machine learning models: logistic regression, bagged random forest and support vector machine. The final best performing support vector machine showed an area under the curve (AUC) of 0.886, with a maximal sensitivity and specificity of 80% and 84.8%. The receiver operating characteristic (ROC) of a similarly trained model using only QTc values, age and gender as inputs, showed an AUC of 0.823, with a maximal sensitivity and specificity of 70.7% and 80%, respectively, to diagnose LQTS.

Conclusion The proposed model resulted in a major rise in sensitivity and a minor rise in specificity compared to the current situation and therefore leads to a decrease in LQTS underdiagnosis. External validation, however, is still necessary to confirm these results.

Contents

Acronyms and parameters

AUC	Area under the curve
aV	Augmented voltage
AV-node	Atrioventricular node
Bagged	Bootstrap aggregated
Ca^{2+}	Calcium
CC	Correlation coefficient
Cl^-	Chloride
EAD	Early afterdepolarization
ECG	Electrocardiogram
ECG_{RMS}	Root mean square ECG
f_s	Sample frequency
HR	Heart rate
ISRoC	Inversed signed radius of curvature
K^+	Potassium
Lasso	Least absolute shrinkage and selection operator
LBB	Left bundle branch
LQT1, LQT2, LQT3	Long QT syndrome type 1, 2 and 3
LQTS	Long QT Syndrome
M-cell	Midmyocardial cell
Na^+	Sodium
PVC	Premature ventricular complex
QTc interval	QT interval corrected for heart rate
RBB	Right bundle branch
RH	R wave heterogeneity
ROC	Receiver operating characteristic
RR	Interval between two R peaks
RSS	Residual sum of squares
SA-node	Sinoatrial node
SM QRS-T angle	Spatial mean QRS-T angle
SP QRS-T angle	Spatial peak QRS-T angle
T_{end}	End of the T wave
T_{pk}	Peak of the T wave
T_{start}	Start of the T wave
TH	T wave heterogeneity
VCG	Vectorcardiogram
YI	Youden's index

Part I

Improving Long QT Syndrome diagnosis using machine learning on ECG characteristics

Chapter 1

Introduction

Congenital long QT syndrome (LQTS) is a genetic disorder affecting cardiac ion channels, increasing the risk of malignant ventricular arrhythmias and sudden cardiac death, see sections 6.3 and 8.1. [1]. Its prevalence is estimated at 1 in 2000 individuals [7]. LQTS is thought to be the cause of death in 20% of sudden unexplained death cases in the young [8]. In rare cases, cardiac arrest can be the first presented symptom of LQTS. [9]. For these reasons, early recognition of LQTS is necessary.

Patients often present themselves with a prolonged QTc interval, which is QT interval corrected for heart rate, as measured in the electrocardiogram (ECG), see section 7.1. However, LQTS can be problematic to diagnose, due to several reasons. Symptoms as syncope and cardiac arrest, which should be prevented in the first place, can occur rarely. A prolonged QTc interval often remains undetected by physicians: only 50% of cardiologists and 40% of noncardiologists are able to correctly identify a lengthened QTc interval. [10] More importantly, there is a significant overlap of QTc intervals between healthy individuals and LQTS patients. Since QTc interval is the first and most widely used diagnostic criterion, this causes severe underdiagnosis of LQTS. A more sensitive tool for LQTS diagnosis in daily clinical practice is required.

Genetic testing provides extra insights in risk stratification and it can confirm the diagnosis of LQTS. However, these tests are unavailable to many centers since this is very costly and specialized care. Also, it can take a significant amount of time before genetic testing results are known. [11]. Additionally, as described in section 8.2, diagnosis cannot be confirmed in an estimated 20% of patients through means of genetic testing, misinterpreting innocent genetic changes as mutations may occur, and a large fraction of genetically identified patients do not show clinical symptoms, due to poor penetrance of the associated genes [12–14]. Moreover, genetic testing is only performed if patients are suspected of having LQTS, e.g. if they have obvious symptoms or if mutations are identified in family members.

Besides a prolonged QTc interval, LQTS is also known to produce aberrant T-waves in the ECG, see section 8.3. An ECG based diagnostic tool could overcome some drawbacks associated with genetic diagnosis of LQTS, such as high costs, time delay and unavailability. This tool would need to be more accurate than the QTc interval. However, an easy-to-use, objective, widespread diagnostic tool for all genotypes of LQTS still lacks.

In light of the above, this study proposes an automated ECG based diagnostic tool for LQTS, including T-wave morphology- and QRS-based features, QTc interval and patient characteristics. A machine learning algorithm will be used so future classification is not prone to time-consuming matters or inter-observer differences.

Chapter 2

Methods

2.1 Study population

ECG recordings were performed in the initial evaluation of individuals >16 years referred to the department of Cardiology and Cardiogenetics of the Academic Medical Centre in Amsterdam, The Netherlands, in the work-up during family screening for LQTS. ECGs were acquired from January 1996 to December 2017. LQTS patients had a confirmed pathogenic mutation in either the KCNQ1, KCNH2 or SCN5A gene resulting in LQTS type 1 (LQT1), type 2 (LQT2) or type 3 (LQT3), respectively (see section 8.2). Controls were genotype-negative family members. Subject gender, age and results of genetic testing were used for this study. Exclusion criteria were the absence of genetic testing results, absence of baseline data, and co-morbidities affecting ventricular depolarization, repolarization and/or the registration of the ECG. Examples included a previous infarction, atrial fibrillation or hypertrophic cardiomyopathy. A waiver was obtained from the local ethical committee, for ethical approval for the conduct of this study.

2.2 Data acquisition

Standard 10-second 12-leads body surface ECGs were obtained at rest. Sample frequencies were 250Hz or 500Hz. ECGs were stored in the MUSE Cardiology Information system (GE Medical System). All further analysis was done using MATLAB R2017a (MathWorks, Natick, MA, USA), as follows: 1) Preprocessing of the signal, to remove noise and obtain a uniform sampling frequency; 2) Construction of an average complex for each lead, to retrieve a smoother signal; 3) Global detection of landmarks, used to calculate morphology features; 4) Calculation of morphology features. For a more detailed overview of the methods, see figure 2.1.

2.3 Preprocessing

R and Q detection filtering. For detection of the R-peak and Q-wave (see section 7.1), a 2nd order high-pass Butterworth filter of 0.5Hz was used. In case of a 500Hz sampling frequency, a 2nd order low-pass Butterworth filter of 125Hz was also applied. Thereafter, an infinite impulse response notch filter of 50Hz was applied. This ECG is referred to as $ECG_{R,Q}$.
T detection filtering. For T-wave detection and analysis, the raw signal was filtered with a 2nd order Butterworth band pass filter with cutoff frequencies 0.5Hz and 40Hz, referred to as ECG_T. [15]
Subsequent preprocessing. The residuals of a median filter with a window of $0.6 \cdot f_s$ (=150 or 300 samples) were subtracted from the individual ECG leads of $ECG_{R,Q}$ and ECG_T to correct for baseline wander. A 1D Fourier upsampling method was used to upsample data to 1000Hz, to assure independence of sampling frequency. Subsequently, nine true unipolar leads were constructed from $ECG_{R,Q}$ and ECG_T, which were used for all further analysis: VR, VL, VF and V1-V6 (see section 7.2). A root-mean-square ECG was constructed from all 9 unipolar leads for both $ECG_{R,Q}$ ($ECG_{RMS^{RQ}}$) and ECG_T (ECG_{RMS^T}).

Data acquisition

10 second, 12-Leads ECGs
f_s = 250 or 500Hz.

Preprocessing

1. 2nd order Butterworth filter.   2. 2nd order IIR notch filter.

3. Baseline deviation removal.   4. Upsampling to 1kHz.

Unipolar lead construction

VR, VL, VF.

ECG_{RMS} construction

Root mean square of 9 unipolar leads.

R peak detection (R_I)

In lead with highest R peaks.
Modified Pan-Tompkins algorithm.

RR_I calculation

Mean value of all RR intervals after removing the 10% higest and lowest values of RRs.

AC_1 construction

For each unipolar lead. Beats are selected from R_I-0.25RR_I to R_I+0.7RR_I.

Beats aligned on R_I. For each sample, the 10% lowest and highest values of the superimposed signals is disregarded. AC_I is the average of the remaining superimposed signals.

AC_F construction

For each unipolar lead

1. Removal of noisy and ectopic beats (CC<0.9 with MB_I).
2. Removal of beats of which RR which deviates >20% from RR_I.
3. For each sample, the 10% lowest and highest values of the <u>remaining</u> superimposed complexes is disregarded. AC_I is the average value of the remaining superimposed signals.

R_{MB} detection

In AC_F of ECG_{RMS}. Modified Pan-Tompkins algorithm.

Q_{MB} detection

Peak of second derivative in ECG_{RMS}.

QT and QTc calculation

Bazett, Fredericia, Framingham, Hodges.

RR_F calculation

Mean value of all remaining RR intervals after removal of the 10% lowest and highest RR intervals.

PCA1 construction

First eigenvector of T wave window in AC_F over all 9 unipolar leads. From R_{AC}+70ms to R_{AC}+0.7RR_F

Morphology feature calculation

T wave detection
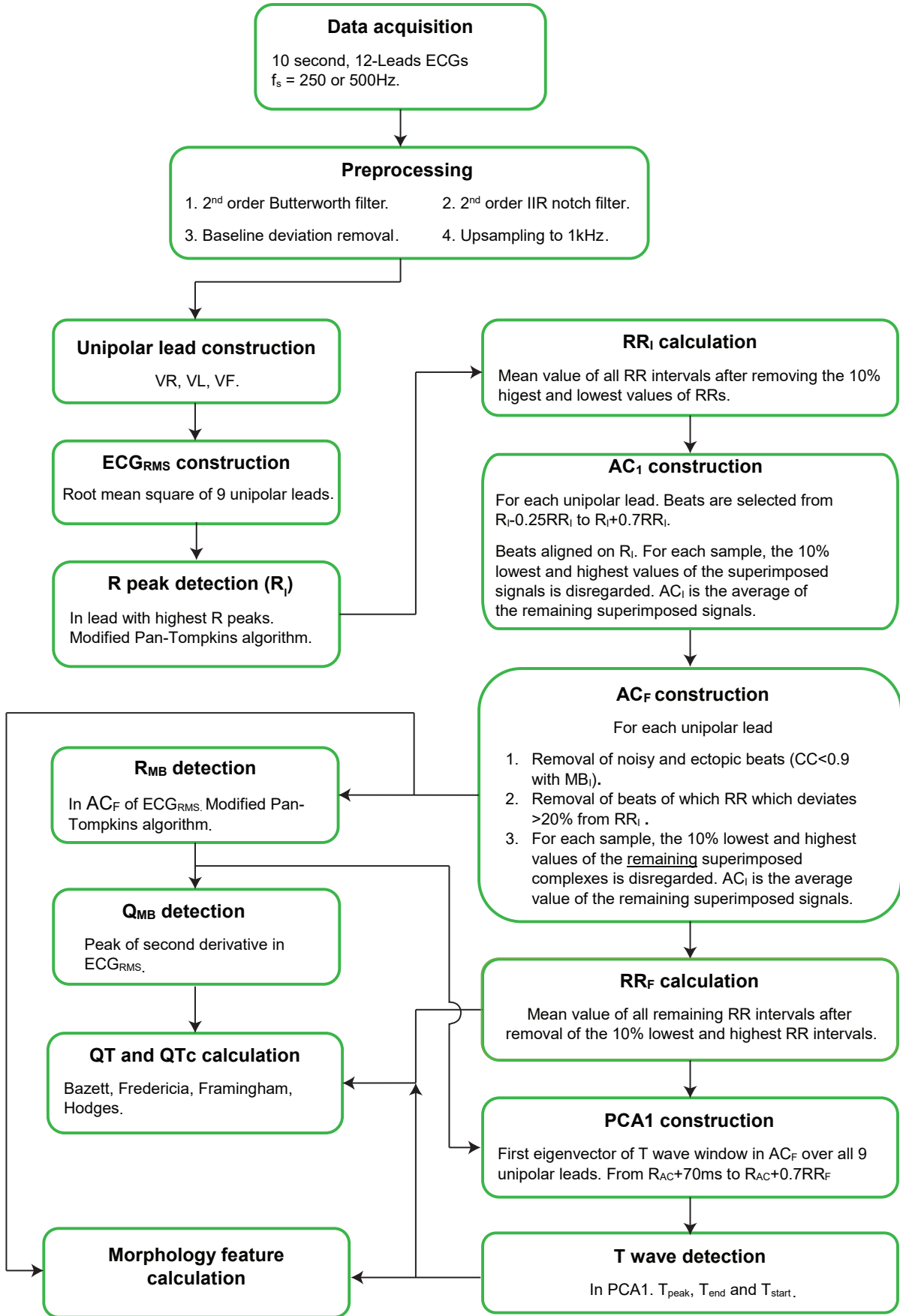
In PCA1. T_{peak}, T_{end} and T_{start}.

Figure 2.1: Diagram of all methods. CC: correlation coefficient.
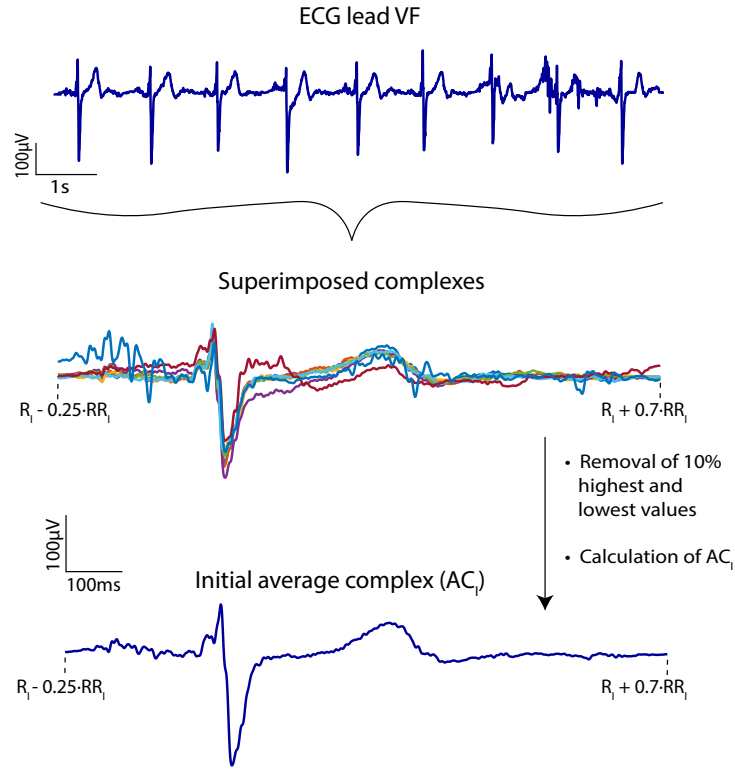
Figure 2.2: Construction of initial average complex

2.4 Average complex construction

Initial average complex construction. R-peaks (R_I, a vector) were detected in the lead with the highest peak amplitudes, using a modified Pan-Tompkins algorithm [16]. R-peaks in the first 0.1 seconds and the last 0.5 seconds of the signal were disregarded. An average complex was constructed for each lead, see figure 2.2. After disregarding the highest and lowest 10% of RR intervals, RR_I was calculated as the average of all remaining RR intervals. Subsequently, complexes were selected from R_I-0.25·RR_I to R_I+0.7·RR_I, for each lead. For each lead, these complexes were superimposed, aligned on the R-peak. The highest and lowest 10% of superimposed signals were disregarded at each sample. An average complex (AC_I) was constructed for each lead, using all remaining complexes.

Final average complex construction. After construction of AC_I, a three-step approach was used to retrieve a smoother signal. An adapted version of the algorithm proposed by Orphanidou et al. [17] was used:

1. If the correlation coefficient (CC) between an individual complexes and AC_I < 0.9, the corresponding complexes was disregarded. If CC ≥ 0.9, the complexes was preserved.

2. If the RR interval of a complex deviated >20% from RR_I, the corresponding complex was disregarded.

3. Subsequently, the number of remaining complexes were investigated.

   (a) If <60% of all complexes in one lead was preserved, the associated lead was not taken into account for analysis.

   (b) If a complex was preserved in <6 leads, the complex was disregarded as a whole.

   (c) If <60% of all complexes in the whole ECG was preserved, the whole ECG was disregarded.

Subsequently, all remaining complexes were aligned on the R-peak again. For each lead, the lowest and highest 10% were disregarded, after which a final average complex (AC_F) was computed. The final average RR interval (RR_F) was calculated in the same manner, using only the RR intervals of preserved complexes. AC_F and RR_F were used for further analysis.
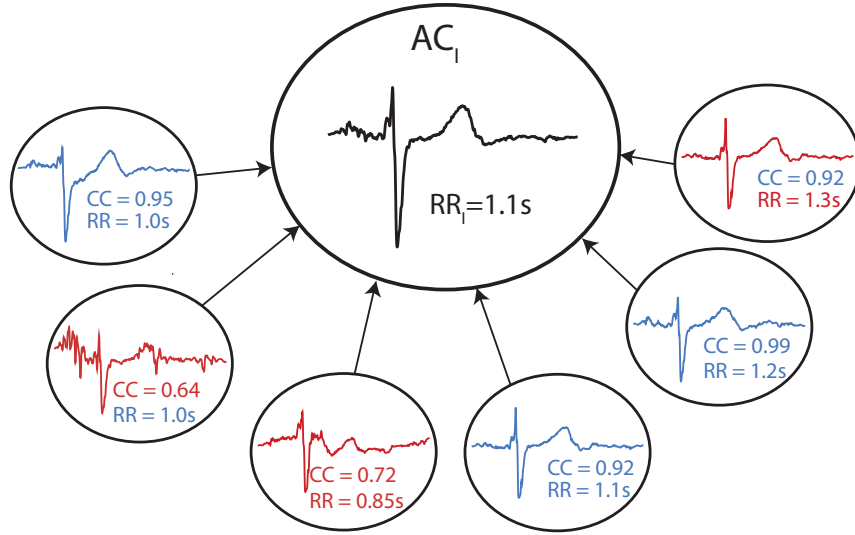
15

Figure 2.3: Step 1 and 2 in constructing AC_2. If the correlation coefficient (CC) of AC_1 and a certain complex was smaller than 0.9, this complex was disregarded. If the RR of a complex deviated more than 20% from the average RR, the complex was disregarded as well. Blue: preserved complexes. Red: disregarded complexes.

2.5 Landmark detection

R-peak detection. A modified Pan-Tompkins algorithm was used to detect the global R-peak of AC_F (R_{AC}), in $ECG_{RMS^{RQ}}$ [16].

Q-wave detection. A simple linear differentiator (see appendix A.1) was used twice to compute the second derivative of AC_F in $ECG_{RMS^{RQ}}$. Subsequently, the most prominent peak of the second derivative in the window from R-100ms to R-20ms was noted as Q.

Principal component analysis. Principal component analysis, using singular value decomposition, was performed over the unipolar leads of AC_F in ECG_T, in the window from R_{AC}+95ms to R_{AC}+0.7·RR_F. The first eigenvector (PCA1) was used for T-wave detection. This signal was filtered with a 2nd order 0.2Hz Butterworth high pass filter, to correct for any deviations from the baseline of 0mV.

T-wave detection. T wave detection was based on the methods by Hermans et al. [18] (see figure 2.4). The peak of the T-wave (T_{pk}) was detected by finding the most prominent peak of PCA1. If the amplitude of T_{pk} was <40mV in ECG_{RMS^T}, the subject was disregarded, to avoid unreliable T-wave start (T_{start}) and T-wave end (T_{end}) detections. Subsequently, the first derivative of PCA1 was calculated, using a linear differentiator (see appendix A.1). T_{start} was detected as the intersection of the baseline with a tangent line drawn at the maximum slope of PCA1, in the window from R+95ms to T_{pk}. T_{end} was detected as the intersection of the baseline with a tangent line drawn at the minimum slope of PCA1, in the window from T_{pk} to T_{pk}+0.3·RR_F. In case of a negative T_{pk}, PCA1 was flipped vertically before T_{start} and T_{end} detection.

2.6 Feature extraction

2.6.1 Included features

Locally determined features. T-wave morphology features were computed within the window between T_{start} and T_{end}, for each lead. Features included: area, absolute area, biphasicness, amplitude, skewness, kurtosis, notch score [4] and asymmetry score [4]. Additionally, QRS amplitude was calculated in each lead [19]. These features were calculated in the unipolar leads VR, VL, VF, and V1 trough V6. Other features, computed in multiple leads, were T-wave heterogeneity from V1-V3 (TH(V1-V3)) and V4-V6 (TH(V4-V6)) [20], and QRS-heterogeneity from V1-V3 (RH(V1-V3)) and V4-V6 (RH(V4-V6)) [20].

Globally determined features. Globally determined features were: T-wave length, T_{pk} to T_{end} interval (TpTe) [19], R to T_{start} interval [3], spatial peak (SP) QRS-T angle [21] and spatial mean (SM) QRS-T
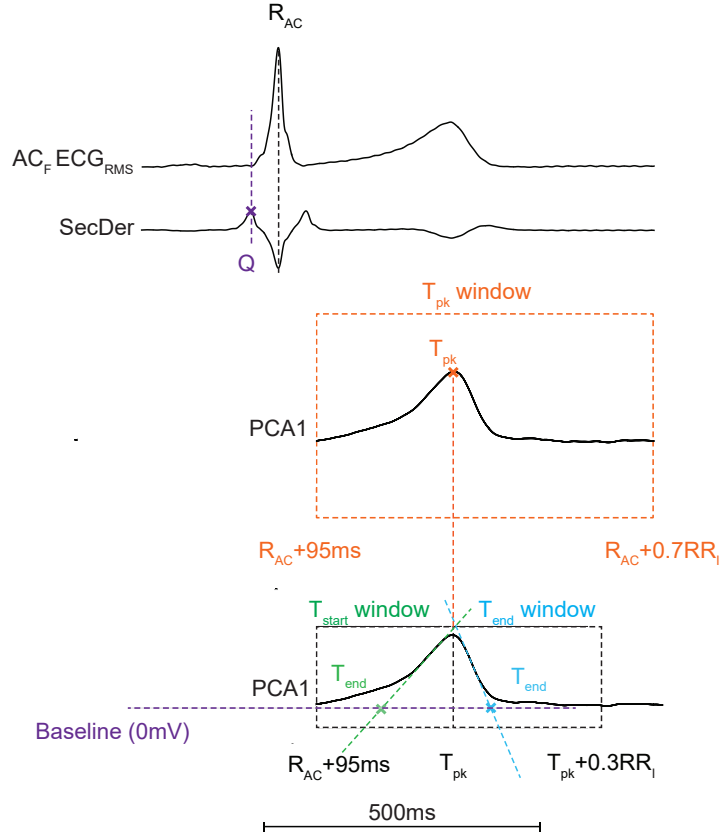
Figure 2.4: Landmark detection. SecDer: second derivative of ECG_{RMS}.

angle [21].

Appendix B shows which leads were used to calculate which features. A total of 9 (leads) · 9 (local morphology features) + 4 (other features) + 5 (globally determined morphology features) was calculated. Additionally, QT interval and corrected QT intervals were calculated, according to table 7.1. A graphical representation of features is displayed in table 2.1, and figures 2.5 and 2.6.

2.6.2 Feature calculation

Calculation of basic features is shown in table 2.1. If a T-wave was biphasic, skewness, kurtosis, notch score and asymmetry score were not calculated. Biphasic was defined as $\frac{Absolute\ area}{area} \leq 0.75$.

Biphasicness The biphasicness of a T-wave was calculated calculated by dividing the absolute value of the area by the absolute area of the T-wave.
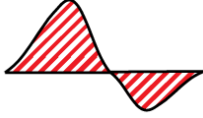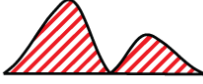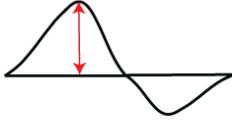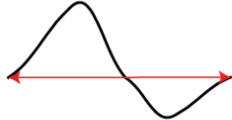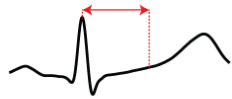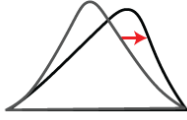
Amplitude To calculate the T-wave amplitude, both the most positive and the most negative peak in the signal were detected. The peak with the largest absolute value was noted as the T-wave amplitude.

Skewness and kurtosis. For the calculation of the T-wave skewness and kurtosis, the T-wave was treated as if it is a probability distribution curve. To this end, T-waves were normalized to have a minimum value of 0 and an area of 1 before calculating the skewness and kurtosis. For more information, see appendices A.2 and A.3.

Notch score and asymmetry score. Notch score and asymmetry score were derived from Andersen et al. [4, 5] However, instead of calculating the notch score only in the first principal component of the signal, a notch was sought for in every lead. A 3rd order Savitzky-Golay filter with a window of 51 samples was used to smooth and differentiate ECG_T.

To calculate the notch score, the Inversed Signed Radius of Curvature (ISRoC) signal was calculated as

Table 2.1: Graphic representation of all features and the formulas used to calculate them. For skewness and kurtosis, a non-biphasic T-wave was used as input, for simplicity.

Feature	Example	Formula
Area		$\sum\limits_{t=T_{start}}^{T_{end}} ECG(t)$
Absolute area		$\left\lvert \sum\limits_{t=T_{start}}^{T_{end}} ECG(t) \right\rvert$
Amplitude		See section 2.6.2
Length		$t(T_{end}) - t(T_{start})$
Time to onset		$t(T_{start}) - t(R)$
Skewness		$\frac{E(ECG-\mu)^3}{\sigma^3}$ (see appendix A.2)
Kurtosis		$\frac{E(x-\mu)^4}{\sigma^4}$ (see appendix A.3)

follows:

$$ISRoC = \frac{\ddot{x}}{(1+\dot{x})^{\frac{3}{2}}} \tag{2.1}$$

where $\ddot{x}$ is the second derivative of the T-wave signal and $\dot{x}$ represents the first derivative of the T-wave signal. Subsequently, a pair of positive and negative values in the ISRoC signal was sought for. In Andersen's study, the height of the positive peak of this pair was noted as the notch score [4]. In contrast, we noted the absolute difference of the pair as the notch score, since the polarity of the ISRoC curve depends on the polarity of the T-wave. If no pair could be found, the notch score was 0. If the amplitude was <40mV, no notch score was calculated, since this would lead to detection of noise as a notch. This method is displayed in figure 2.5.

To calculate the asymmetry score, the first derivative of the ECG was divided into two segments: segment 1, from T_{start} to T_{pk}, and segment 2, from T_{pk} to T_{end}. Both segments were scaled between 0 and 1. Segment 2 was flipped over both the x and y-axis. The shortest segment of the two was supplemented with zeros at the start of the segment, until both segments were equally long. Subsequently, the asymmetry score was calculated (see figure 2.5).

QRS-T angle. A vectorcardiogram (VCG) was produced (see section 7.3), in which the orthogonal projections of the ECG on the X-, Y- and Z-axis are displayed. From this signal, the spatial peak QRS-T angle and the spatial mean QRS-T angle were computed. The start of the Q-wave is needed to calculate QRS-T angles. If no Q_{AC} was detected, Q_{AC} was denoted as R-50ms.
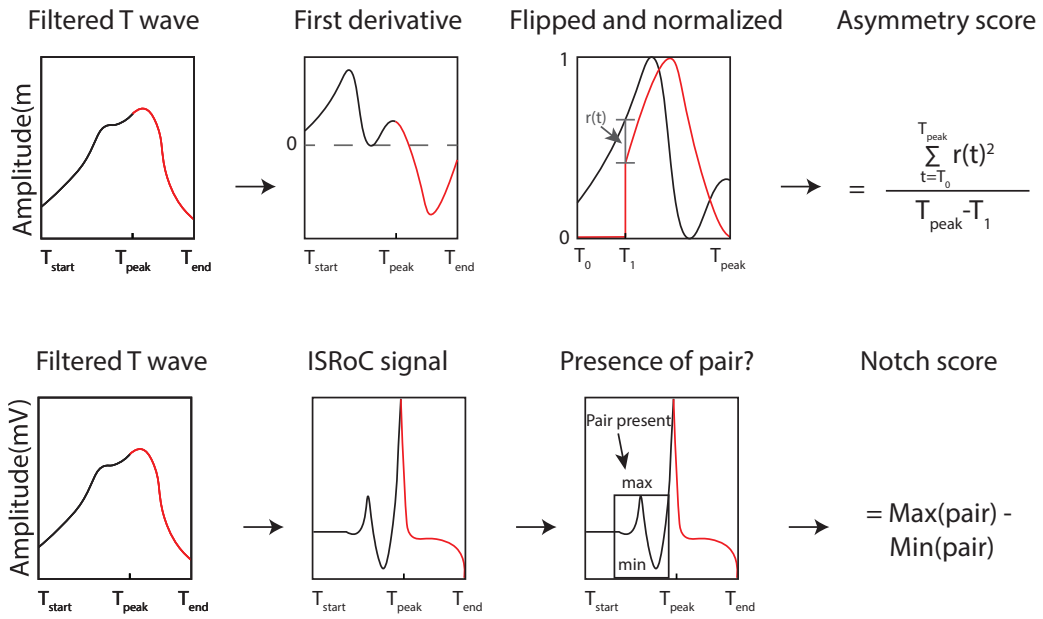
Figure 2.5: Asymmetry score and notch score. **Asymmetry score**: 1) the first and second segment are separated at T_{pk}. 2) The second segment is flipped over the x- and y-axis. The shortest segment is supplemented with zeros at the start of the signal, until both segments are equally long. 3) The asymmetry score is calculated. T_1 is the difference in length between both segments. **Notch score**: a notch creates an up-down pair in the ISRoC (Inversed Signed Radius of Curvature) signal. The absolute difference between the amplitudes of this pair in the ISRoC signal is noted as the notch score. Adapted from [4].

Figure 2.6: Spatial QRS-T angle as calculated from the VCG. Red: QRS-loop. Blue: T-loop. The smallest angle between the vectors at maximal T-wave magnitude and maximal QRS magnitude is called the 'spatial peak QRS-T loop'.

Spatial peak QRS-T angle. The spatial peak QRS-T angle is the smallest angle between the vector at maximal T-wave magnitude in the VCG and the vector at maximal QRS complex magnitude in the VCG, see figure 2.6. For more information, see appendix A.4.
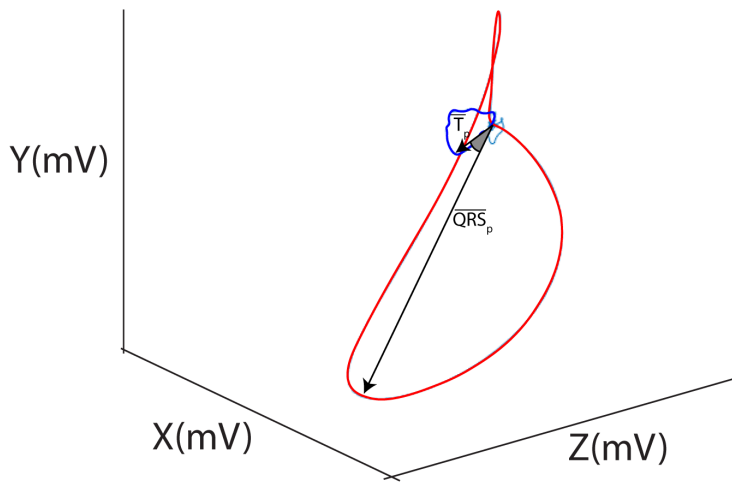
Spatial mean QRS-T angle. The spatial mean QRS-T angle is the smallest angle between the mean vector of the T-wave and the mean vector of the QRS-complex in the VCG (see figure 2.6). For more information, see appendix A.5.

R-peak and T-wave interlead heterogeneity. To calculate R-peak heterogeneity, the QRS-complex was selected from Q to R+50ms. If no Q was detected, Q was denoted as R-50ms. For T-wave heterogeneity, the T-wave was selected from T_{start} to T_{end}. Two groups of precordial leads were investigated: V1-V3 and V4-V6. Subsequently, the maximum value of the square root of the variance around the average of the three leads was calculated [8, 22], as follows:

$$H = \max(\sqrt{var(X)}) \tag{2.2}$$

in which X is an n-by-3 matrix consisting of three single-lead vectors of length n. The resulting variance is a vector of length n.

2.7 Machine learning models

2.7.1 Inputs and models

A publicly available package, 'glmnet', was used for model training and testing. [23] Subjects were classified as LQTS or healthy by a machine learning trained classification model, based on multiple inputs, see section 2.6.1. Missing observations were replaced by random values within mean±sd (standard deviation) for the corresponding feature in the corresponding lead. To average out the effect of using random values, different random values were used three times. The performance of each single model was averaged over these three random values. Two models were trained: a baseline model with only subject age, gender, and all QT(c) values (see table 7.1) as inputs, and a final model with all morphology features as additional inputs. The baseline model is trained and tested similarly as the final model. The baseline model was used to assess the optimal classification using currently available clinical inputs. The difference between the baseline model and the final model demonstrates the added diagnostic value of morphology features.

2.7.2 Final model

The final product of this thesis is a machine learning model which can diagnose LQTS. For this model, the entire study population is used as a training set. However, if the entire study population is used as a training set, the model cannot be tested, hence the performance of the model is not known. For this reason, to assess the performance of the final model, 100 models are trained and tested on a different randomized training and testing set, see figure 2.7. Training sets consisted of a randomly chosen subset containing 90% of all subjects, testing sets consisted of the remaining 10% of all subjects. The mean performance of these 100 models is the expected performance of the final model. [24]

2.7.3 Assessment of final model performance

Feature selection. Each of the 100 models was trained and tested separately. Features with the highest discriminative performance were selected by means of elastic net regularization, combined with maximum likelihood estimation in a logistic regression model (see sections 9.1.1 and 9.2). The elastic net mixing parameter γ was varied from 0 to 1, with steps of 0.2. For each of these 6 γ values, the elastic net tuning parameter λ was decreased in 100 steps. For each λ, the cross-validated error was noted. The cross-validated error was found using 10-fold cross-validation, see section 9.1.2. In other words: $10 \cdot 6 \cdot 100 = 6000$ models were trained for each randomized training set, to find the optimal set of discriminative features for this training set. Subsequently, for each γ, the coefficients (β) resulting from the cross-validation were noted at λ_{min} (at the minimal cross-validated error) and λ_{1SE} (one standard error away from λ_{min}).
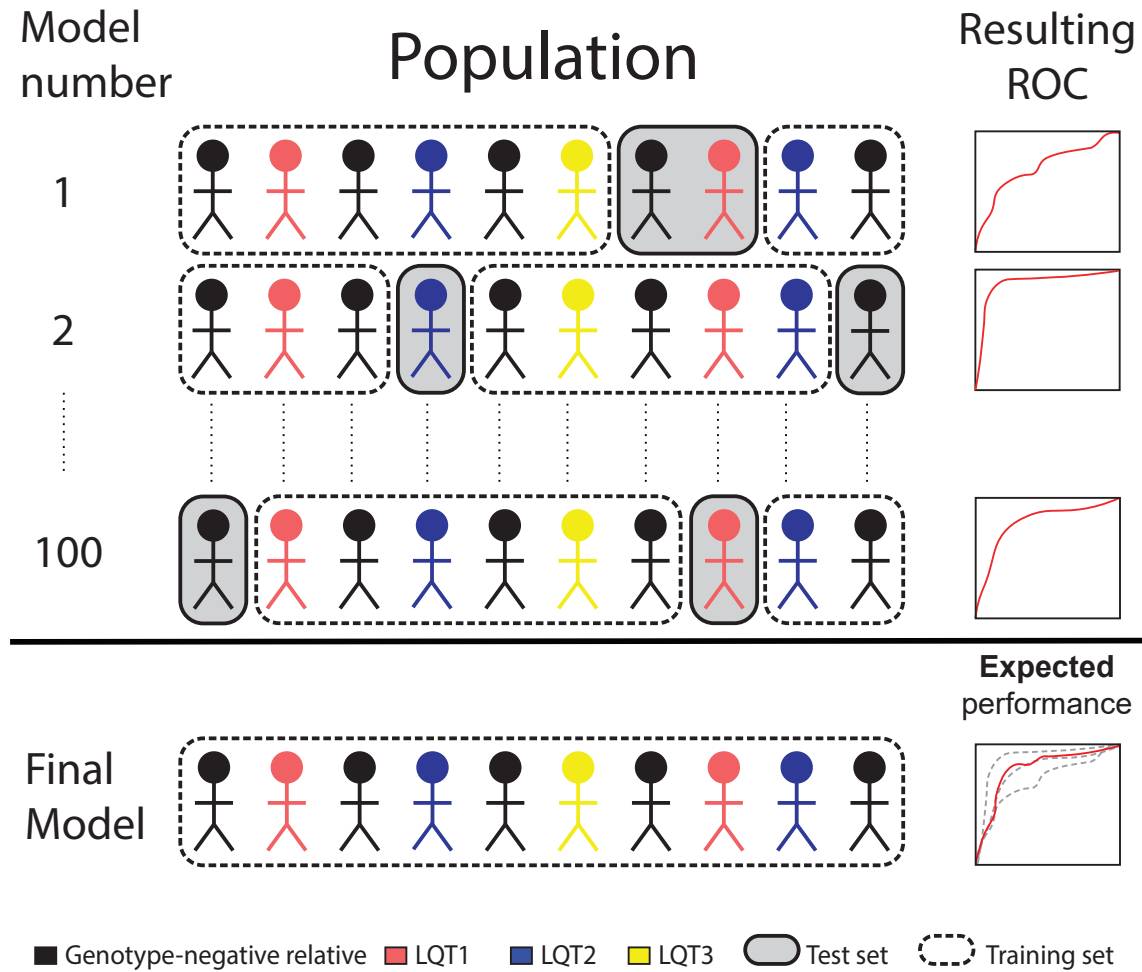
Figure 2.7: Construction and assessment of final model. For simplicity, only 10 subjects are drawn. Training and test set are randomized 100 times. The expected performance of the final model is the mean of the 100 models. In this example, the training set consisted of 80% of all subjects, and the testing set consisted of 20% of all subjects. Note that although on average subjects were chosen 80 times for training and 20 times for testing, this does not hold for all individual subjects, because of randomization.

Model training and testing. After feature selection, for each of the 100 randomizations, 3 different types of models were trained: logistic regression, a bagged random forest, and a support vector machine, see section 9.2. These models were trained for features selected at λ_{min} and λ_{1SE}, for each of the 6 γ values. In other words: $2 \cdot 3 \cdot 6 \cdot 100 = 3600$ models were trained after feature selection. Finally, the model with the best mean discriminative performance over all 100 randomizations (e.g. a logistic regression model with γ=0.6 at λ_{min}) was chosen for final model training on all subjects. For this final model, the included features and β resulting from feature selection (see section 9.1.1) are reported.

Chapter 3

Results

3.1 Study population

A total of 1123 subjects with unique ECGs were recorded. The exclusion flowchart is shown in figure 3.1. The remaining cohort consisted of 699 subjects. Baseline data are shown in table 3.1. Age between LQTs patients and genotype-negative relatives was significantly different, with LQTS patients being 4 years older, on average. Gender did not differ significantly between LQTS patients and genotype-negative relatives. Age and gender between LQTS genotypes did not differ significantly either.

Table 3.1: Baseline data of research population. Relative: genotype-negative relatives. Note that LQT1-LQT3 are genotypes of LQTS. †: two-tailed student's t-test. ‡: one-way anova. $^{\mathrm{II}}$: chi-squared test.

Characteristic	Relatives (n=349)	LQTS (n=350)	p-value	LQT1 (n=135)	LQT2 (n=162)	LQT3 (n=53)	p-value
Age	45±15	41±15	3E-4†	41±14	42±15	40±15	0.78‡
Male gender	164 (47%)	148 (42%)	0.21$^{\mathrm{II}}$	55 (41%)	72 (44%)	21 (40%)	0.74$^{\mathrm{II}}$

3.2 QTc interval

All QT(c) interval measurements were manually screened by an experienced researcher (TD). Since the algorithm was not able to find the end of the T-wave fully reliable in case of biphasic or flat T-waves, these were adjusted manually for 30 subjects.

Using only the QTc interval as an input in a linear regression model to diagnose LQTS, the receiver operating characteristic (ROC) shows an area under the curve (AUC) of 0.781-0.812, depending on the method of QTc calculation. The corresponding ROCs are shown in figure 3.2. The best performing QTc calculation method was Hodges', with an AUC of 0.812. The maximal Youden's index (YI$_{max}$) (=sensitivity+specificity-1) is 0.482. The corresponding QTc interval threshold was 435ms, irrespective of gender. The corresponding sensitivity and specificity were 0.660 and 0.822, respectively.

Table 3.2: Performance of different sorts of QT(c) to diagnose LQTS.

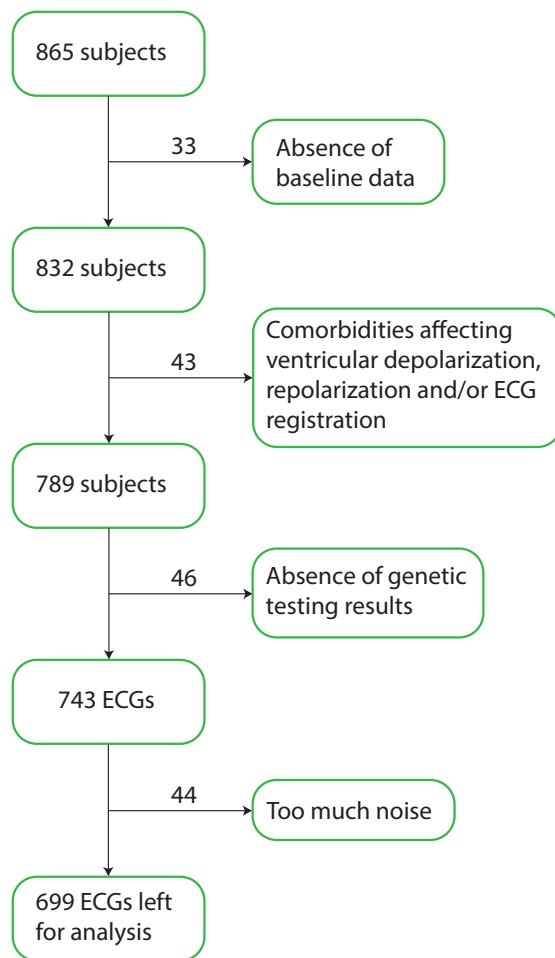| Interval	AUC	YI$_{max}$	Sensitivity	Specificity	Threshold at YI$_{max}$ (ms)
QT	0.781	0.443	0.569	0.874	433
QTc Bazett	0.782	0.440	0.590	0.851	450
QTc Fridericia	0.806	0.482	0.671	0.811	435
QTc Framingham	0.803	0.477	0.666	0.811	434
QTc Hodges	0.812	0.482	0.660	0.822	435

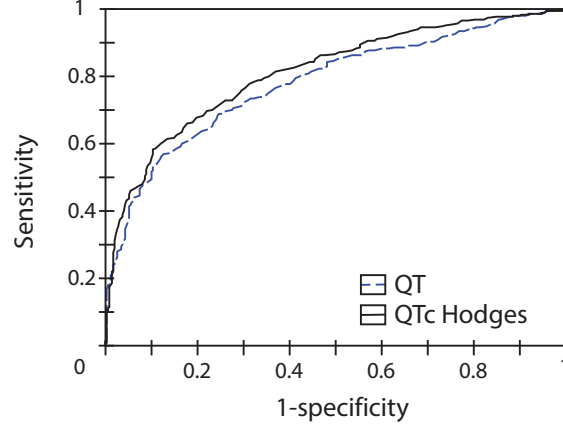Figure 3.1: Exclusion flowchart. 'Too much noise' defined by section 2.4.

Figure 3.2: ROCs of different kinds of QT and QTc Hodges.

3.3 Machine learning

Baseline model. Regarding the baseline model, the model with the best performance was the logistic regression model, combined with a tuning parameter γ of 1. Results are displayed in table 3.3 and figure 3.3.

Final model performance. For each model, γ leading to the best performance is reported. The ROCs of the different types of models are shown in figure 3.3. γ, AUC, and YI are shown in table 3.3. Figure 3.4 displays the performance of 1000 support vector machines with $\gamma = 0.4$ and $\lambda = \lambda_{1SE}$ on all subjects.

Table 3.3: Performance of models with best tuning parameter γ. Criterion: elastic net mixing parameter λ_{min} or λ_{1SE}.

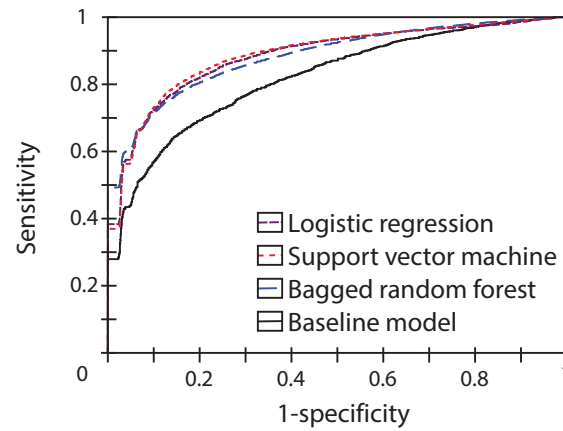| Model	Best γ	Criterion	AUC	95% CI	YI_{max}	Sensitivity	Specificity
Baseline model	1	λ_{min}	0.823	0.812-0.833	0.507	0.707	0.800
Logistic regression	1	λ_{min}	0.884	0.876-0.892	0.634	0.769	0.865
Bagged random forest	1	λ_{min}	0.882	0.874-0.890	0.623	0.748	0.875
Support vector machine	0.4	λ_{1SE}	0.886	0.878-0.895	0.648	0.800	0.848

Figure 3.3: Mean interpolated ROC of all used models and QTc Hodges. The baseline model was a logistic regression model, with only QT(c) values, age and gender as inputs. The performance was averaged over all 100 models.
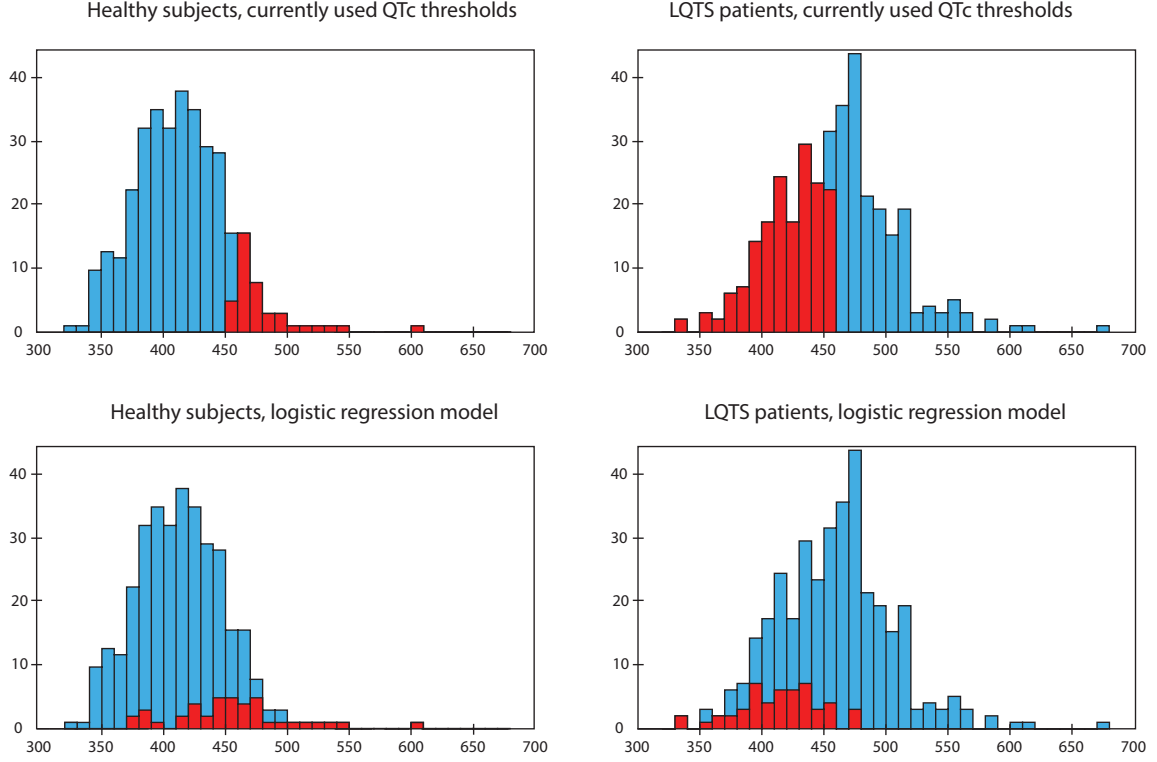
Figure 3.4: Histograms displaying the performance of the currently used QTc thresholds (450ms for men and 460ms for women, using QTc Bazett), and our support vector machines. Blue: correctly classified subjects. Red: incorrectly classified subjects. To produce this figure, 1000 models were used on their own testing set (see figure 2.7). All models operated at YI_{max}. The average prediction of all 1000 models per subject is shown here.

Included features in final model. Since a support vector machine with $\gamma=0.4$ performed the best, this model was trained on all data. For $\lambda = \lambda_{1se}$, the 10 most important features were: Asymmetry in VR and V3, skewness in VR, VF, V5 and VL, T_{pk} to T_{end} interval, kurtosis in VL and V3, and biphasicness in VR. For a list of all included parameters and their coefficients, see appendix D.

Chapter 4

Discussion

4.1 Machine learning model

When using a machine learning approach combined ECG-based morphology features and subject characteristics, the best mean model AUC was 0.886, as opposed to a maximal AUC of 0.812 when only using QTc thresholding. At a maximal Youden's index, this means a rise of 14.0% sensitivity and 2.4% in specificity is achieved, compared to the QTc thresholding, leading to their respective values of 80% and 84.8%. Compared to the baseline model, the added value of T-wave morphology shows an increase in AUC, sensitivity and specificity of 0.063, 9.3% and 4.8%, respectively. As visible in figure 3.4, the added value of our method is found in healthy subjects with a Bazett's QTc>450 (or 460, for females) and patients with a Bazett's QTc≤450 (or 460, for females). The proposed model could decrease LQTS underdiagnosis if implemented in clinical practice, since a major rise in sensitivity and a minor rise in specificity are achieved when diagnosing LQTS, compared to QTc thresholding.

In our study, our results were compared the genetic testing results. As stated in chapter 1, a large fraction of genetically identified patients do not show clinical symptoms, due to poor penetrance of the associated genes [12–14]. Genotype-negative relatives can have a QTc of >500ms, while LQTS patients can have a QTc of <400ms. Therefore, it could be more relevant to compare our results to the symptomaticity of patients instead of to genetic testing results, as some other studies have attempted [25]. Unfortunately, data on symptomaticity were not available for our study.

Our study contained LQT1, LQT2 and LQT3 patients and compared them to genotype-negative relatives, which resembles reality. In other studies where LQTS was diagnosed through the ECG by Andersen et al. and Chorin et al., healthy volunteers were compared to LQTS patients [4, 6]. This could lead to an increase in classification accuracy compared to the real-life scenario. Chorin et al. found a sensitivity and specificity were 82% and 77% when classifying LQTS at rest, respectively [6]. Andersen et al. did not include any LQT1 or LQT3 patients, which can also have a positive effect on classification accuracy. They reported a 90% sensitivity and a 95% specificity when classifying LQT2 patients and healthy controls [4]. Both studies did not use a training set and a test set to classify subjects, which could increase prediction accuracy as well, considering that the optimal cutoff values can be found for the entire research population.

As stated in chapter 1, diagnosis cannot be confirmed in an estimated 20% of patients through means of genetic testing. Although our study might have contained positively tested patients without symptoms, we are certain about the genetic diagnosis of all subjects. Since genotype-negative subjects in our study were relatives of included LQTS patients, no false-negative controls were included, since a subject's relative should have the same mutation as the subject himself.

A standard 12-leads ECG is typically a part of standard care in cardiology, which could be a major advantage for implementation in clinical practice. Other attempts to diagnose LQTS without the use of genetic testing have also been made, for example by using an epinephrine test or by investigating the mechano-electrical window in echocardiography. [25, 26] The main disadvantage of most studies is that additional interventions need to be performed.

Our model selected a combination of skewness, kurtosis, and asymmetry, biphasicness and T_{peak} to T_{end} interval as the 10 most discriminative features. According to articles of Chorin et al. and Yan et al., the

time to onset, length, amplitude, area, notch should be more discriminative to diagnose LQTS. [3, 6] However, their research focused on manually assessing pseudo-ECGs or pairs of leads. Our method automatically combines the outcomes of all leads separately, which could have led to a different set of discriminative features. Our findings are however in agreement with the findings of Andersen et al., in general. They found that notch score, asymmetry score and flatness score (=1-kurtosis) discriminate LQT2 patients from healthy subjects. [4] However, as stated in section 4.3, notch score was not used as an input in our research.

Recommendations

It needs to be emphasized that in our study, the ratio of patients to controls was approximately 1:1. This is not representative for the 1:2000 prevalence of LQTS in the general population [7]. For this reason, it should be considered to choose a different cutoff value with a higher specificity. Otherwise, an overdiagnosis of LQTS - or at least a high number of referrals for LQTS - could occur.

The models produced in our study have only been tested internally, on a known cohort. Even though the average performance of 100 models was calculated to assess the potential performance of a final model trained on all data, this final proposed model has not been tested on an external cohort yet. Additionally, it is a topic of discussion what the final model should be. Two options are available: 1) to use an ensemble of 100 models and use their mean prediction for classification, or 2) to train one final model on all data. The first option poses problems in terms of interpretability. For the second option, it is less clear how this final model would perform on an external validation set. Both options should be considered and tested on an external validation set.

Even though the quality of the 10-second 12-leads ECGs in this research most often did not allow to analyze individual beats, it would be eligible to do so. In our study, these filtering methods and the calculation of an average complex were necessary to obtain a smooth signal. However, in some cases, LQTS has been reported to produce T-wave alternans, which are beat-to-beat variations in T-wave morphology [27, 28]. ECGs in this study were part of standard clinical routine. At time of registration, it was not yet known that these ECGs would be used for automated analysis. For future implementation, we would advise to give additional care to the quality of the ECG when recording.

Our study focused on an automated method to diagnose LQTS. Some improvements need to be made regarding the detection of biphasic and flat T-waves in our algorithm, and external validation is still needed. However, given the performance and the ease-of-use of the model, it does show potential be implemented in clinical care as a first diagnostic tool to indicate the potential diagnosis LQTS.

4.2 QTc threshold

When using only the individual QT or QTc intervals to classify subjects as healthy or LQTS, Hodges' QTc yielded the best results in terms of ROC AUC: 0.812. Sensitivity and specificity at YI_{max} were 66% and 82.2%, respectively. The cutoff value was 435ms. Hodges' QTc was followed by Fridericia (AUC=0.806), Framingham (AUC=0.782), and Bazett (AUC=0.781), consecutively.

Our study used a global principal component analysis based detection. This method was chosen because in LQTS, QT intervals can vary greatly between leads (see section 8.1). However, in clinical practice, the QT interval is measured in lead II, V5 or V6. In 30 patients (4%), T-wave end could not be detected correctly by the used algorithm. Reasons included biphasic T-waves, flat T-waves or high interlead variability in Tend. Since our study aimed to develop an observer-independent algorithm, this is one drawback which should be resolved in the future.

Currently, the most universally used correction method is Bazett's method, ever since its first introduction in 1920 [29]. However, as multiple studies have indicated before, Bazett's correction tends to undercorrect for heart rates (HR) smaller than 60 beats per minute (bpm), and overcorrect for HR > 60 bpm [30–32].

Wong et al. compared six different QT interval correction methods in 1179 caucasian athletes with or without bradycardia, concluding that Fridericia's QTc is the best method for clinical interpretation of QT interval in this study population [30]. Luo et al. concluded that Hodges' QTc shows the least correlation with HR in 10303 healthy caucasians, suggesting that Hodges' method is the best method of correction [31]. This shows that different methods of QT correction could be more appropriate than

the widespread Bazett's QTc. However, both studies assessed a different goal than our study: they both aimed to find the QTc method with the least correlation with HR. When aiming for the most widespread applicability of QT correction and an LQTS diagnosis rate as accurate as possible, we advise to use Hodges' QTc combined with a cutoff value of 435ms, irrespective of gender. In our research, this led to a sensitivity of 66% at 82.2% specificity, as opposed to a sensitivity of 51.3% and a specificity of 89.8% when using Bazett's QTc with the currently maintained thresholds of 450ms for men and 460ms for women [33].

4.3   Notch score

The methods of Andersen et al. were used to calculate a notch score [4, 5]. Validation of this notch score yielded very poor agreement between the calculated notch score by the algorithm and the manual notch assessments by two of our researchers: YI was <0.1. For this reason, the calculated notch score was not included as a predictor to the machine learning models.

In our study, other filtering methods were used than the study performed by Andersen et al. [4,5]. Even though an average complex was constructed in our study and multiple filtering methods were applied to calculate and validate the notch score, the signal was not smooth enough to correctly determine a notch score, since the second derivative is very sensitive to noise.

In the study performed by Andersen et al., a low-pass Kaiser finite impulse response filter with a cutoff frequency of 20Hz was applied to the the first principal component of the (estimated) ST-T segment of the vectorcardiogram. However, the presence of a notch in this principal component was not validated with the presence of a notch in the actual ECG in their study. Additionally, the effect of 20Hz filtering on the presence of a notch was not investigated. Hence, it is unclear whether the notch calculation as determined by Andersen et al. represents a notch in the actual electrocardiogram. For this reason, our research focused on validation of their notch score-algorithm. Still, it would be desirable to validate the presence of a notch in a quantitative manner for LQTS diagnosis, since the presence of a notch has been widely reported in LQT2. [3, 4]

Chapter 5

Conclusion

In this thesis, a machine learning-based diagnostic tool for long QT syndrome was developed. Receiver operating characteristics showed a substantial increase in sensitivity and specificity compared to the current situation, in which a certain QTc threshold as a diagnostic criterion. The best mean model area under the curve of the receiver operating characteristic was 0.897, as opposed to a maximal area under the curve of 0.815 when only using QTc thresholding. At a maximal Youden's index, this means that a rise of 21.2% sensitivity is exchanged for a fall of 1.8% in specificity, leading to their respective values of 79% and 87.5%. While considering the current proposed tool still requires some adjustments and validation, it does show potential for implementation in clinical practice. If the current method of QTc thresholding would be replaced by the diagnostic tool as developed in this study, it could potentially decrease long QT syndrome underdiagnosis.

Part II

Background

Chapter 6

Anatomy and physiology

6.1 The heart as a pump

The human heart is divided into four chambers: the left ventricle, the left atrium, the right ventricle and the right atrium. Blood enters the heart at the atria, and leaves the heart through the ventricles. The right side of the heart receives blood from the body and pumps it through the pulmonary circulation, which carries blood to the lungs and returns it to the left side of the heart. In the lungs, carbon dioxide is removed from the blood, while oxygen is added to the blood. The left side of the heart pumps blood through the systemic circulation, which delivers oxygen and nutrients to the body. From those tissues, carbon dioxide and other waste products are carried back to the right side of the heart.

The contraction of the heart is performed in *cardiac cycles*. Throughout one cardiac cycle, the following steps occur:

1. Rapid ventricular filling. The majority of inflow into the ventricles occurs passively during this rapid ventricular filling phase. Also, the atria fill with blood, coming from:

   - The inferior and superior vena cava and the coronary sinus, for the right atrium.
   - The pulmonary veins, for the left atrium.

2. Atrial contraction. The atria contract, forcing the last 20% of blood into the ventricles.

3. Isovolumetric contraction. The ventricles contract, while no ejection takes place yet. Pressure builds up in the ventricles.

4. Ventricular ejection. The ventricles contract further, pumping the blood into:

   - The aorta, from the left ventricle.
   - The pulmonary arteries, from the right ventricle.

Once contraction has finished and the blood has left the ventricles, the cardiac cycle starts over again. A schematic representation of the basic anatomy of the human heart is displayed in figure 6.1(a).

6.2 The heart as an electrical system

An extensive electrical system is needed to generate and conduct electrical impulses which cause cardiomyocytes (cardiac muscle cells) to contract. At the start of each cardiac cycle, an electrical impulse is generated at the sinoatrial node (SA-node), which is situated in the superior wall of the right atrium. After its generation, this impulse spreads rapidly over the atria, making the atria contract. The electrical impulse takes one of the internodal pathways to travel to the atrioventricular node (AV-node). After a delay at the AV-node, the electrical impulse travels further through the AV-bundle, which is situated inferiorly to the AV-node. Subsequently, the impulse spreads through the left and right bundle branch (LBB and RBB). The bundle branches terminate in Purkinje fibers, which are located in the inner ventricular walls. The Purkinje fibers pass the electrical activity to the myocardial tissue, making the ventricles contract simultaneously. A schematic representation of the basic anatomy and physiology of the electrical system of the heart is displayed in figure 6.1(b).
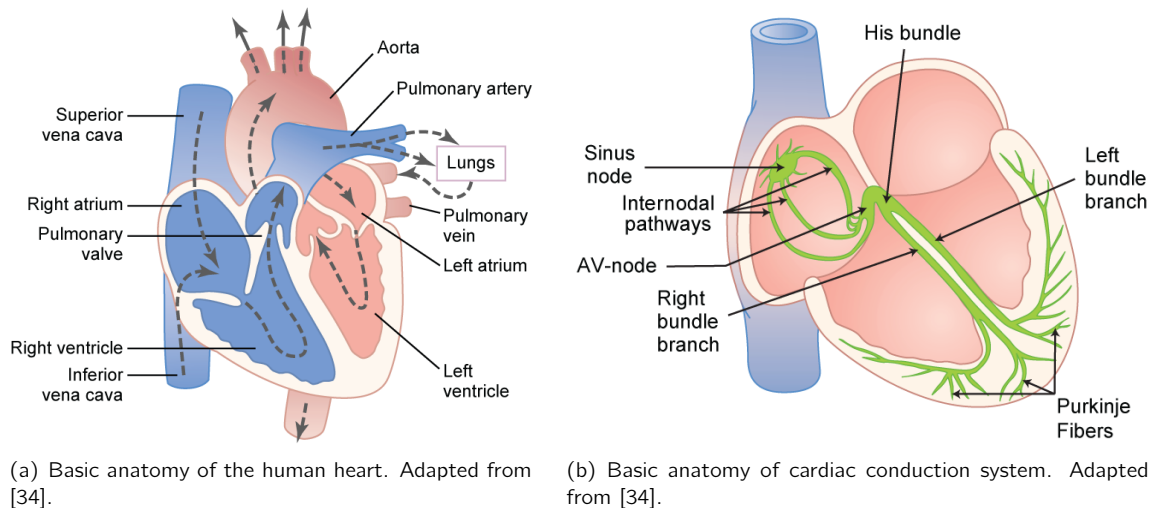
(a) Basic anatomy of the human heart. Adapted from [34].

(b) Basic anatomy of cardiac conduction system. Adapted from [34].

Figure 6.1: Anatomy of the human heart.

6.3 The cardiac action potential

The generation and conduction of electrical impulses occur on a cellular level. The SA-node is the primary natural pacemaker of the heart, consisting of a specialized group of cells that can generate electrical impulses at regulated intervals. These electrical impulses are generated by ion flows which produce *action potentials*. These action potentials spread throughout the heart, following the route described in 6.2. The cells on the inside of the heart, called *cardiomyocytes*, are coupled through electrical bridges, called *gap junctions*. A cardiac action potential consists of five phases. These phases differ somewhat between cardiac cells, but as an example, the events within a ventricular cardiomyocyte will be presented, since this cell type is most relevant considering the scope of this study. Figure 6.2 displays the five phases of an action potential phases of a ventricular cardiomyocyte, which are as follows [35]:

Phase 4. In this phase, the cardiomyocyte cell membrane is at rest. Among all ions, the cell is most permeable to free potassium (K^+) ions, which explains why the resting cell membrain potential is roughly -85mV. Ionic pumps maintain this cell potential.

Phase 0. During this phase, a rapid, positive change in voltage across the cell membrane occurs. The positive change in voltage is called *depolarisation*. This depolarization originates from activation of sodium (Na^+) channels, which increases the flow of Na^+ ions over the cell membrane. Sodium channels are activated by the arrival of an action potential of a neighboring cardiomyocyte through a gap junction. If the arrival of this action potential increases the cell membrane potential to a certain threshold value, the sodium channels will open, depolarizing the cardiomyocyte.

Phase 1. During this phase, Na^+ channels are rapidly deactivated, drastically reducing the movement of Na^+ ions into the cell. At the same time, K^+ channels open for a very brief amount of time, leaving K^+ ions out of the cell, and decreasing the membrane potential slightly more.

Phase 2. In this phase, the membrane potential remains almost constant, while the membrane begins to *repolarize* very slowly. This means a negative change in membrane potential occurs. This allows K^+ ions to leave the cell. However, calcium (Ca^{2+}) ions and chloride (Cl^-) ions flow into the cell, which almost leaves the membrane potential constant. The Ca^{2+} inflow does three things:

1. Binding to calcium channels on the sacroplasmatic reticulum, opening them. This allows Ca^{2+} to flow out of the sacroplasmatic reticulum, letting the cardiomyocyte contract.

2. Indirectly mediating (Cl^-) ion outflow, which opposes the voltage change caused by the K^+ outflow.

3. (Indirectly) increasing activity of the sodium-potassium pump.

As a result of these ion movements, the net membrane potential will barely change.

Phase 3. In the final phase of the cardiac action potential, rapid repolarization takes place. Three vital steps occur:

1. The Ca^{2+} ion flow is stopped.

2. More K^+ ions leave the cell. Consequently, a net outward positive current occurs, which causes the cell to repolarize.

3. Ionic pumps restore ion concentrations back to the pre-action potential state, meaning that intra-cellular Ca^{2+} ions are pumped out of the cell, which in turn means that the cardiomyocyte stops contracting. Overall, there is a net outward positive current during phase 3, producing a negative change in membrane potential.
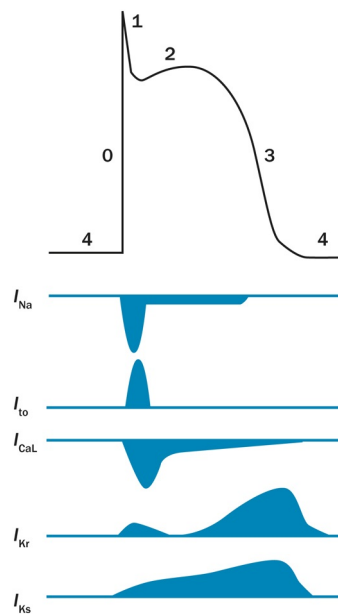

Figure 6.2: The five phases of a ventricular action potential. Ion currents of Na^+, K^+ and Ca^{2+} ions is denoted. Reproduced from [36].

Chapter 7

Electrocardiography

## 7.1 Physiology behind electrocardiography

Action potentials occur in all cardiomyocytes, which means this process occurs in endocardial, epicardial and midmyocardial cells. Activation starts at the endocardium, since the Purkinje fibers terminate in this area. Thereafter, the midmyocardial layer and the epicardium are activated. During an electrocardiogram (ECG) recording, the net sum of these voltages is measured in different directions, typically defined as described in section 7.2. This is shown in figure 7.1(a). By subtracting the epicardial signal from the endocardial signal, the pseudo-ECG arises. This pseudo-ECG is a transmural ECG through the ventricular myocardium, measured at one location. During cardiac activity, this can be performed in multiple directions, resulting in an ECG which displays the net activation vector in each direction. The ECG has several distinguished waveforms, which all correspond to a certain electrical activation of cardiomyocytes:

1. **The P-wave** corresponds to the depolarization of both atria.

2. **The QRS-complex** corresponds to the depolarization of both ventricles.

3. **The T-wave** corresponds to the repolarization of both ventricles.

An example of these characteristic waveforms is displayed in figure 7.1(b).
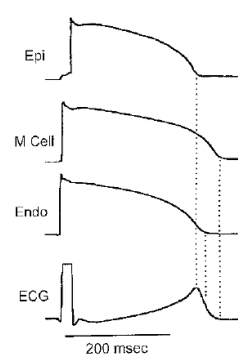
(a) Cellular basis of the ECG. The relationship between the epicardial, M-cell and endocardial action potential produces the net activation through the myocardium. Reproduced from [3].
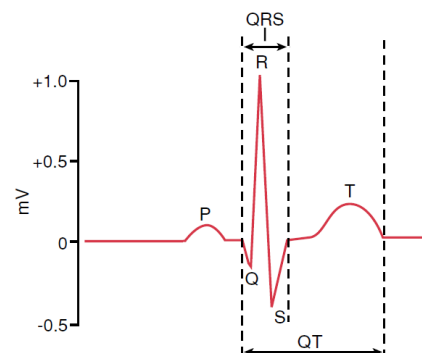
(b) The waveforms of an ECG, all corresponding to certain electrical activity of the heart. P: atrial depolarization. QRS: ventricular depolarization. T: ventricular repolarization. Adapted from [37].

Figure 7.1: Cellular basis of the ECG and ECG waveforms

Table 7.1: Different methods to correct QT interval for heart rate. RR: interval between current and previous R-peak. HR: Heart rate in beats per minute.

Correction method	Formula
Bazett	$QTc_B = \frac{QT}{\sqrt{RR}}$
Fridericia	$QTc_{Fc} = \frac{QT}{\sqrt[3]{RR}}$
Framingham	$QTc_{Fh} = QT + 0.154 \cdot (1 - RR)$
Hodges	$QTc_H + 1.75 \cdot (HR - 60)$

The QT-interval as shown in figure 7.1(b) is often corrected for heart rate, since repolarization tends to shorten with increasing heart rate. Multiple correction formulas have been developed: Bazett's, Fridericia's, Framingham's, or Hodges' method (see table 7.1). Bazett's method is applied the most.

7.2   Technical background

An ECG can be acquired by placing electrodes on the body surface. The net sum of all electrical activity in volts in a certain direction in the heart is summed and measured in the ECG by measuring potential differences. The ECG is the most used clinical tool to assess cardiac electrical activity, to identify abnormalities in heart rhythm, conduction or cardiac ischemia. The ECG electrodes to measure potential differences can be placed in different configurations. The most common way of acquiring an ECG is by using the limb leads, in combination with the augmented voltage leads and the precordial leads.

7.2.1   Limb leads

The limb leads, are the original ECG leads as defined by Einthoven, only using three electrodes. One electrode is placed on the right arm (r), one electrode is placed on the left arm (l) and one is placed on the left leg (f). Additionally, a ground electrode is always placed on the right leg as a reference, to prevent power line noise from interfering with the small biopotential signals of interest. For simplicity, this electrode is not shown in figures. The three potential differences are called *limb leads* I, II and III [38]:

$$\text{I} = \phi_l - \phi_r \tag{7.1}$$
$$\text{II} = \phi_f - \phi_r \tag{7.2}$$
$$\text{III} = \phi_f - \phi_l \tag{7.3}$$

where ϕ_r is the potential measured at r, ϕ_l is the potential measured at l, and ϕ_f is the potential measured at f. These leads are called *bipolar*, since they directly measure potential differences between the electrodes. Einthoven's leads are displayed in figure 7.2(a).

7.2.2   Augmented voltage leads

The augmented voltage leads, as first described by Goldberger, are constructed by using the limb leads. They contain no information that was not already present in the limb leads, but the six signals are easier to interpret by visual inspection. Augmented voltage leads are termed *unipolar leads* because a single positive electrode is referenced against a combination of the other limb electrodes. Augmented voltage (aV)-leads are constructed as follows [38]:

$$aVR = \phi_r - \frac{1}{2} \cdot (\phi_l + \phi_f) = -\frac{1}{2} \cdot (I + II) \tag{7.4}$$

$$aVL = \phi_l - \frac{1}{2} \cdot (\phi_r + \phi_f) = \frac{1}{2} \cdot (I - II) \tag{7.5}$$

$$aVF = \phi_f - \frac{1}{2} \cdot (\phi_r + \phi_l) = \frac{1}{2} \cdot (II + III) \tag{7.6}$$

The augmented voltage leads are displayed in figure 7.2(a).
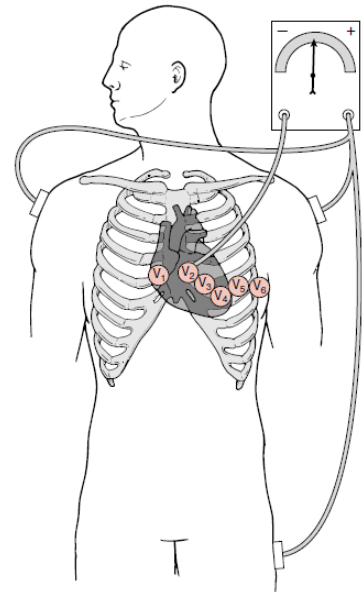
7.2.3 Precordial leads

The precordial leads (also referred to as Wilson's leads), are unipolar leads as well. The potential difference is measured between each precordial electrode and the average of ϕ_r, ϕ_l and ϕ_f. These potential differences are called V_1 through V_6. The precordial electrodes all have their own specific anatomical reference on which they should be placed (see figure 7.2(b)):

- V_1 on the fourth intercostal space, just right to the sternum

- V_2: on the fourth intercostal space, just left to the sternum

- V_3: halfway between V_2 and V_4

- V_4: on the left fifth intercostal space, midclavicular line

- V_5: horizontal to V_4, anterior axillary line

- V_6: horizontal to V_5, mid-axillary line
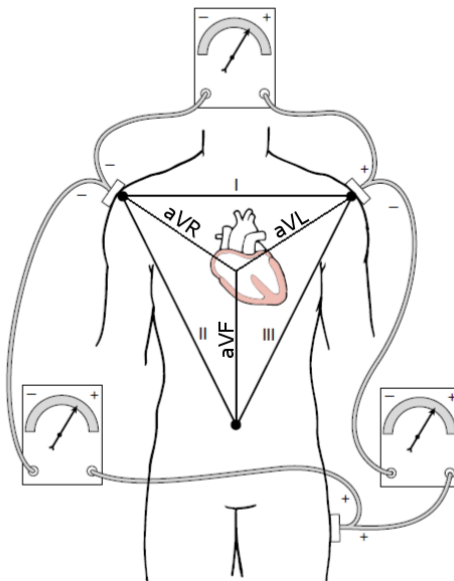
(a) Limb leads and augmented voltage leads. Adapted from [39].

(b) Precordial leads. Note that the three limb leads are used as a reference. Reproduced from [39].

Figure 7.2: ECG leads

7.2.4 True unipolar limb leads

Mathematically, augmented limb leads are scaled true unipolar limb leads (hence the name 'augmented'). For example, the unipolar VR lead is calculated as follows:

$$\text{VR} = \phi_r - \phi_{WCT} = \phi_r - \frac{\phi_r + \phi_l + \phi_f}{3} = \frac{2 \cdot \phi_r - \phi_f - \phi_l}{3} = \frac{2}{3} \cdot \text{aVR} \tag{7.7}$$

where ϕ_{WCT} is the Wilson central terminal (the electrical center of the frontal plane electrodes). The augmented limb lead avR is calculated by:

$$\text{avR} = \phi_r - \frac{\phi_l + \phi_f}{2} = \frac{2 \cdot \phi_r - \phi_f - \phi_l}{2} \tag{7.8}$$

So, the true unipolar limb leads are calculated by scaling the augmented voltage variants with two thirds [18]:

$$\text{VR} = \frac{2}{3} \cdot \frac{2 \cdot \phi_r - \phi_f - \phi_l}{2} = \frac{2 \cdot \phi_r - \phi_f - \phi_l}{3} = \frac{2}{3} \cdot \text{aVR} \tag{7.9}$$

7.3 Vectorcardiography

As explained in section 7.2, the net electrical activity of the heart in certain directions is displayed in the ECG. Instead of measuring the net electrical activity in 12 predefined directions like in a 12-leads ECG, the vectorcardiogram (VCG) displays a reconstruction of the absolute electrical activity in a 3D space, defined by orthogonal X-, Y-, and Z-axes, containing all electrical information. The three leads are represented by right-left axis, head-to-feet axis and front-back (anteroposterior) axis. [40] The VCG can be reconstructed from the original 12-lead ECG, see equations (7.10) to (7.12).

$$X = -(-0.172 \cdot V_1 - 0.074 \cdot V_2 + 0.122 \cdot V_3 + 0.231 \cdot V_4 + 0.239 \cdot V_5 + 0.194 \cdot V_6 + 0.156 \cdot I - 0.010 \cdot II) \tag{7.10}$$

$$Y = (0.057 \cdot V_1 - 0.019 \cdot V_2 - 0.106 \cdot V_3 - 0.022 \cdot V_4 + 0.041 \cdot V_5 + 0.048 \cdot V_6 - 0.227 \cdot I + 0.887 \cdot II) \tag{7.11}$$

$$Z = -(-0.229 \cdot V_1 - 0.310 \cdot V_2 - 0.246 \cdot V_3 - 0.063 \cdot V_4 + 0.055 \cdot V_5 + 0.108 \cdot V_6 + 0.022 \cdot I + 0.102 \cdot II) \tag{7.12}$$

where X is the electrical activity in mV on the X-axis, Y the activity on the Y-axis and Z the activity on the Z-axis. An example VCG is shown in figure 7.3.
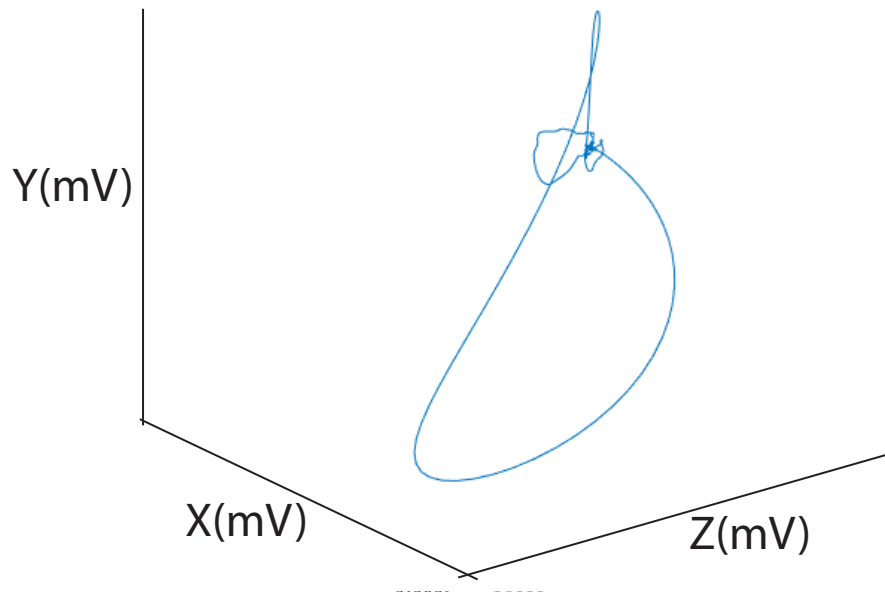
Figure 7.3: Example vectorcardiogram. Each point in space represents the 3D electrical activity at a certain point in time.

Chapter 8

Long QT syndrome

8.1 Pathophysiology

The congenital long QT syndrome is a genetic disorder in which cardiac ion channels are affected. A variety of ion channels can be affected, as described in section 8.2. This leads to action potential abnormalities, particularly in the repolarization phase of the action potential. Repolarization is generally prolonged in LQTS. Since the *penetrance* of affected genes can be very low in LQTS, the degree of repolarization prolongation can vary widely [14]. This means that individuals can carry an affected gene, without having a prolonged repolarization phase of the action potential. This poses the main challenge in diagnosing long QT syndrome.

Since not all cardiomyocytes are affected in the same manner, heterogeneity of repolarization is increased, called *increased dispersion of repolarization*. As described in section 8.3, this can lead to notable changes in the repolarization phase of the ECG. In LQTS, as a consequence of prolonged repolarization and dispersion of repolarization, an extra phenomenon can occur. Groups of clustered cells can produce an extra depolarization before the action potential is finished. This phenomenon is called *early afterdepolarization (EAD)* (see figure 8.1). An EAD is defined as 'a slowing or reversal of normal repolarization during phase 2 or phase 3 of the action potential' [41]. An EAD occurs when the net outward current required for repolarization is compromized. In general, midmyocardial cells (M-cells) are more prone to EADs than endocardial or epicardial cells [41]. Although the existence of the M-cell is still under debate, this discussion is beyond the scope of this thesis. The main point is that repolarization is dispersed along the transmural axis of the myocardium. Impaired ion currents in LQTS can lead to an additional depolarization in this phase, even though the cell has not repolarized fully yet.
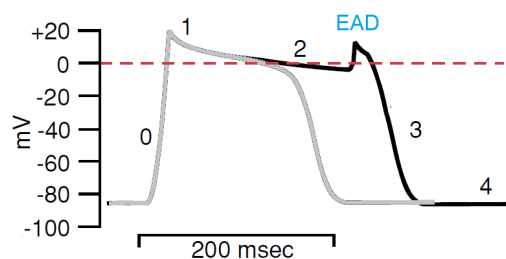
Figure 8.1: Example of an early afterdepolarization. Note the increased repolarization length. Adapted from [39].

Under the right conditions, a group of local EADs can lead to depolarization of adjacent cells, which leads to activation of both ventricles: a premature ventricular complex (PVC). In principle, this is a harmless mechanism without adverse consequences. However, if an EAD occurs under the right circumstances (i.e. the right place and time, it can create so-called *re-entry mechanisms*, see figures 8.2 and 8.3. If re-entry occurs, not all cells are repolarized at the time of EAD formation: these cells are still *refractory*. EAD-forming cells can create a current which starts circling around the refractory cells.

By the time the cycle is almost completed, the previously refractory cells are fully repolarized, making them excitable again. These can in turn make the original cells depolarize again, creating a continuous loop, see figure 8.3. This phenomenon is known as a potentially lethal type of ventricular tachycardia, called *Torsades de Pointes*, see figure 8.2. [41]
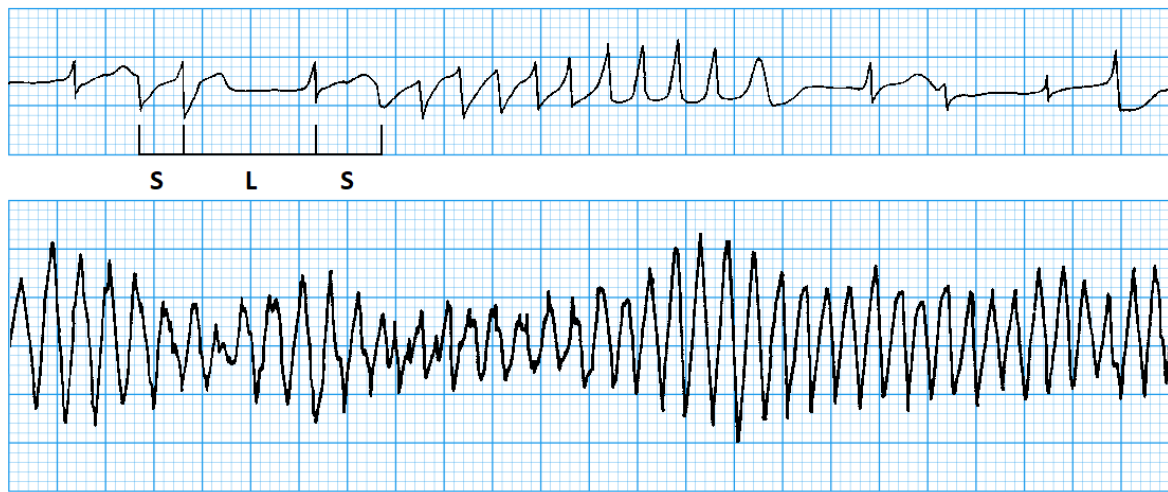

Figure 8.2: **Top**: typical initiation of Torsades de Pointes. A PVC (which follows an **S**) is typically followed by a compensatory pause (denoted **L**). Consequently, the normal beat which follows this complex contains an increased QT interval, or repolarization duration. If a PVC falls during this QT interval, Torsades de Pointes can be initiated, since the ventricles were only partially repolarized. **Bottom**: sustained Torsades de Pointes. Adapted from [42].
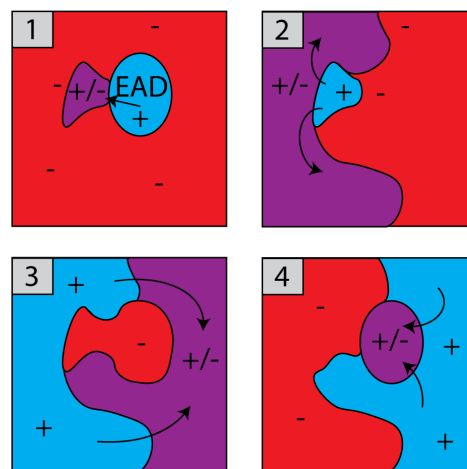
Figure 8.3: Example of a reentry circuit following from an early afterdepolarization (EAD). Each image represents tissue in the ventricular myocardium. Red (-): refractory tissue. Blue (+): depolarizing tissue. Purple (+/-): relative refractory tissue, which is almost ready to be depolarized. The arrows show the direction of the depolarization front.

8.2 Genotypes

Up until now, at least 16 different gene abnormalities have been associated with LQTS. The most frequently occurring types - LQT1, LQT2 and LQT3 - account for 75% of the affected population. Another 5% is affected by the remaining genetic mutations, while 20% of affected patients remain with a clinical diagnosis without any identified mutation so far. [12] In case of an absent (identified) mutation, patients are diagnosed based on clinical presentation, such as family history and presentation of symptoms [33]. Genetic mutations causing LQTS all affect different proteins, which in turn cause

Table 8.1: The tree most frequently occurring types of LQTS with corresponding gene defects and affected ion currents.

Genotype	Gene defect	Affected ion current
LQT1	KCNE1	IKs (slow rectifier K^+ current)
LQT2	KCNE2	IKr (fast K^+ current)
LQT3	SCN5A	INa (Na^+ current)

a variety of cardiac ion channels to be affected. Table 8.1 shows the affected ion channels and genes in LQT1, LQT2 and LQT3. K^+, Na^+ or Ca^{2+} channels can be affected, depending on the altered gene. Although each individual genotype presents itself with its own symptoms and ECG alterations, all mutations lead to repolarization deficits, potentially leading to very dangerous cardiac arrhythmias like Torsades de Pointes. [43]

8.3   T-wave morphology

As a direct consequence of the dispersion of repolarization described in section 8.1, morphology alterations can arise in the repolarization phase (the T-wave) of LQTS patients (see figure 7.1(b)). As a consequence of ionic pump defects, the ECG can show abnormal T-wave morphologies, depending on the particular genotype. An example of altered T-wave morphologies depending on ionic pump alterations in a pseudo ECG is shown in figure 8.4.

The end of the M-cell action potential corresponds with the end of the T-wave in the pseudo-ECG and the end of the epicardial action potential corresponds with the peak of the T-wave in the pseudo-ECG. In all phenotypes of LQTS, the repolarization duration is increased. In LQT1, M-cell repolarization and the epicardial repolarization show a deflection, which makes its corresponding T-wave broad and long. In LQT2, a low-amplitude T-wave appears, because of the small difference between the three signals during the whole repolarization tract. In LQT3, the epicardial repolarization stops relatively early compared to the other two signals, which makes the corresponding T-wave asymmetric. Also, the difference between the three signals remains close to zero for a long time, which makes the T-waves late-onset. [6, 44, 45] Dispersion of repolarization is prone to inducing Torsades de Pointes, see section 8.1. [3, 46].
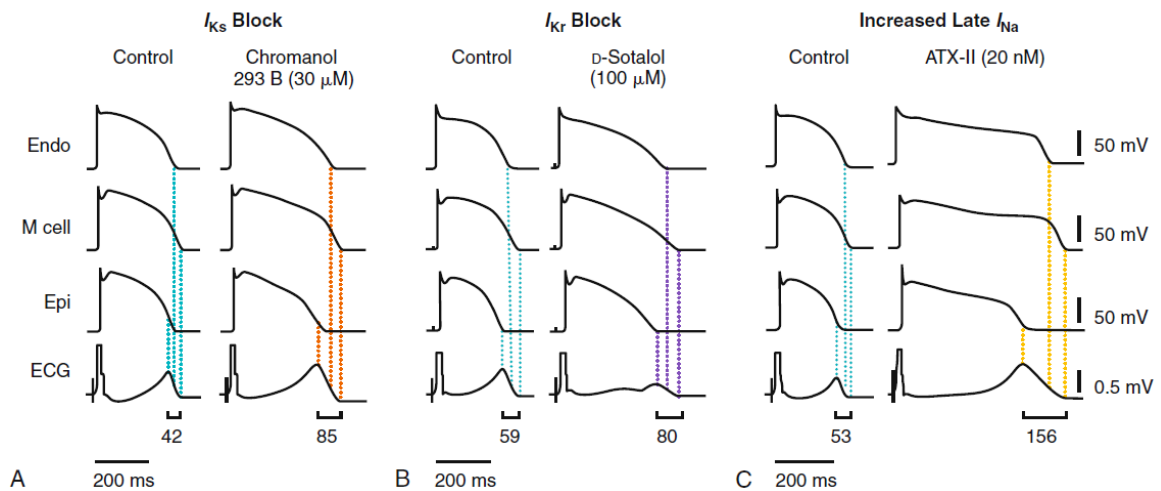
Figure 8.4: Abnormal T-waves in pseudo-ECGs due to induced ion channel defects. The combination of endocardial, M-cell and epicardial signals produces aberrant T-waves. Although the existence of the M-cell is still under debate, the main point is that repolarization is dispersed along the transmural axis of the myocardium. **A**: impaired slowly-activating delayed rectifier K^+ current (as in LQT1) **B**: impaired rapidly-activating delayed rectifier K^+ current (as in LQT2). **C**: impaired Na^+ current (as in LQT3). Reproduced from [46].

Chapter 9

Machine learning

9.1 Feature selection

9.1.1 Elastic net regularization

The following equation describes multivariate linear regression for one subject, which is used to develop a predictive model:

$$f_j(x_j) = \alpha + \beta \cdot x_j + \varepsilon_j \tag{9.1}$$

where $f_j(x_j)$ is the regressand for a particular subject, α is the intercept with the y-axis (which is the response for this subject when all regressors are zero), β is a row vector containing all coefficients, x_j is a column vector containing all regressors (variables used to predict y_j), and ε_j is the error term for this particular subject. A linear fit can be obtained by minimizing the residual sum of squares (RSS):

$$RSS = \sum_{j=1}^{K} (\varepsilon_j)^2 = \sum_{j=1}^{K} (y_j - (\alpha + \beta \cdot x_j))^2 \tag{9.2}$$

where K is the number of subjects and y_j is the measured value of the variable to predict, for each subject. RSS can be minimized by adapting β, to fit the model. This method can work in simple cases. However, in more complex or elaborate cases (e.g. when K is relatively large), an ordinary least squares solution does not predict future data accurately and is difficult to interpret. [24] To increase model interpretability and prediction accuracy, ridge regression or LASSO (least absolute shrinkage and selection operator) can be applied. Ridge regression shrinks the values of β towards zero. However, none of the elements of β is exactly equal to zero, leading to difficulties in model interpretability. On the other hand, LASSO shrinks as much elements of β as possible to zero, while removing correlated features. However, in case of high correlations between predictors, prediction performance of LASSO is dominated by ridge regression. The *elastic net* method combines advantages of both ridge regression and LASSO. [47] The elastic net method selects the most relevant features to include in the model. The elastic net equation describes this method mathematically:

$$\min_{(\alpha,\beta)\in\mathbb{R}^{N+1}} R_\lambda(\alpha,\beta) = \min_{(\alpha,\beta)\in\mathbb{R}^{N+1}} \left(\sum_{j=1}^{K} \left(y_j - (\alpha + \beta \cdot x_j) \right)^2 + \lambda \left(\gamma \cdot ||\beta||_{\ell_1} + \frac{1}{2}(1-\gamma) \cdot ||\beta||_{\ell_2}^2 \right) \right) \tag{9.3}$$

$$\text{where} \qquad ||\beta||_{\ell_1} = \sum_{i=1}^{M} |\beta_i| \tag{9.4}$$

$$||\beta||_{\ell_2}^2 = \sum_{i=1}^{M} \beta_i^2 \tag{9.5}$$

where λ is the elastic net tuning parameter, $||\beta||_{\ell_1}$ is the ℓ_1 penalty leading to feature selection shrinkage as in LASSO, $||\beta||_{\ell_2}^2$ is the ℓ_2 penalty leading to coefficient shrinkage as in ridge regression, γ is the elastic net mixing parameter, and M the number of parameters [23]. Note that this equation is an altered version of equation (9.2). Both γ and λ are varied to find a solution R_λ as small as possible. γ

can be varied from 0 to 1, while λ can be varied from 0 to ∞, in principle. To find the best γ and λ, the cross-validated error is investigated, which is described in section 9.1.2. In our study, the probability of a subject having LQTS is modeled, which means that the outcome is dichotomous, meaning that our subjects are either healthy or diseased. For this reason, logistic regression (see section 9.2.1) is used instead of linear regression. Hence, equation (9.3) should be converted from linear to logistic regression. First, since the probability of a subject belonging to a certain class is modeled, the natural logarithm of the maximum likelihood function (equation (9.10)) from logistic regression is calculated:

$$\ln\left(\prod_{j=1}^{K}\left[\frac{1}{1+e^{-(\alpha+\beta\cdot x_j)}}\right]^{y_j}\cdot\left[1-\frac{1}{1+e^{-(\alpha+\beta\cdot x_j)}}\right]^{1-y_j}\right)=\sum_{j=1}^{K}\left[-\ln\left(1+e^{-(\alpha+\beta\cdot x_j)}\right)+\beta\cdot y_j\cdot x_j\right] \quad (9.6)$$

where y_i is the subject's real disease state. This formulation is also called the *log-likelihood*. The full derivation of this formula can be found in appendix C. Subsequently, the linear regression part in equation (9.3) is replaced by this log-likelihood. The RSS needed to be minimized, to minimize the prediction error in linear regression (see equation eq:RSS). In contrast, the likelihood of correct classification should be maximized. For this reason, the elastic net penalty is subtracted from the maximized log-likelihood.

$$\max_{\alpha,\beta\in\mathbb{R}^{N+1}}\left[\frac{1}{K}\sum_{j=1}^{K}y_j\cdot(\alpha+\beta x_j)-\ln\left(1+e^{-(\alpha+\beta\cdot x_j)}\right)-\lambda\left(\frac{1}{2}(1-\gamma)\cdot||\beta||^2+\gamma\cdot||\beta||\right)\right] \quad (9.7)$$

By maximizing the log-likelihood for the optimal γ and λ, the optimal combination regressors to include can be found.

9.1.2 Cross-validation

Elastic net regularization (see section 9.1.1) can help to select the best subset of regressors for a model, making use of cross-validation. λ and γ are varied iteratively to find a solution R_λ as small as possible (see equation (9.3) and (9.6)). Let γ be a vector of N values; in our study, $N = 6$. Let λ be a vector of Q values, in our study, $K = 100$. Remember that each combination of λ and γ leads to its own β, leading to its own subset selection. Since α is the intercept (which is the response when all elements of β are equal to 0), α does not change if λ or γ is varied. Cross-validation is used for $Q \cdot N$ times, to assess the performance of each β. Eventually, the β with the best performance will lead to the optimal set of features to include in the model.

Cross-validation internally tests each subset of β. In this study, 10-fold cross-validation is used, meaning that 10 differently fitted logistic regression models are produced for each combination of γ_i and λ_j. Data are separated into 10 equally large groups. For each iteration out of 10, a model is trained on 9 of these groups, and tested on the last one, see figure 9.1. Each time a new model is tested, the prediction error ε_m is noted. Summed over all iterations, the total cross-validation error for this (γ_i, λ_j) is equal to:

$$E(\gamma_i, \lambda_j) = \sum_{m=1}^{10}\varepsilon_m(\gamma_i, \lambda_j) \quad (9.8)$$

The cross-validation error is the prediction error on the testing set, summed over all 10 iterations. Since we are only interested in subsets with a low cross-validation error, the included coefficients β are noted at two instances:

1. Value of the elastic net tuning parameter λ (λ_{min}) which leads to the minimal cross-validation error.

2. Value of λ where the cross-validation error is 1 standard error removed from the minimal cross-validation error(λ_{1SE}).

λ_{1SE} works approximately as good for model training as λ_{min}, while including less predictors. This leads to a simpler model, which leads to better interpretability. For this reason, the included β at λ_{1SE} are investigated. [24]
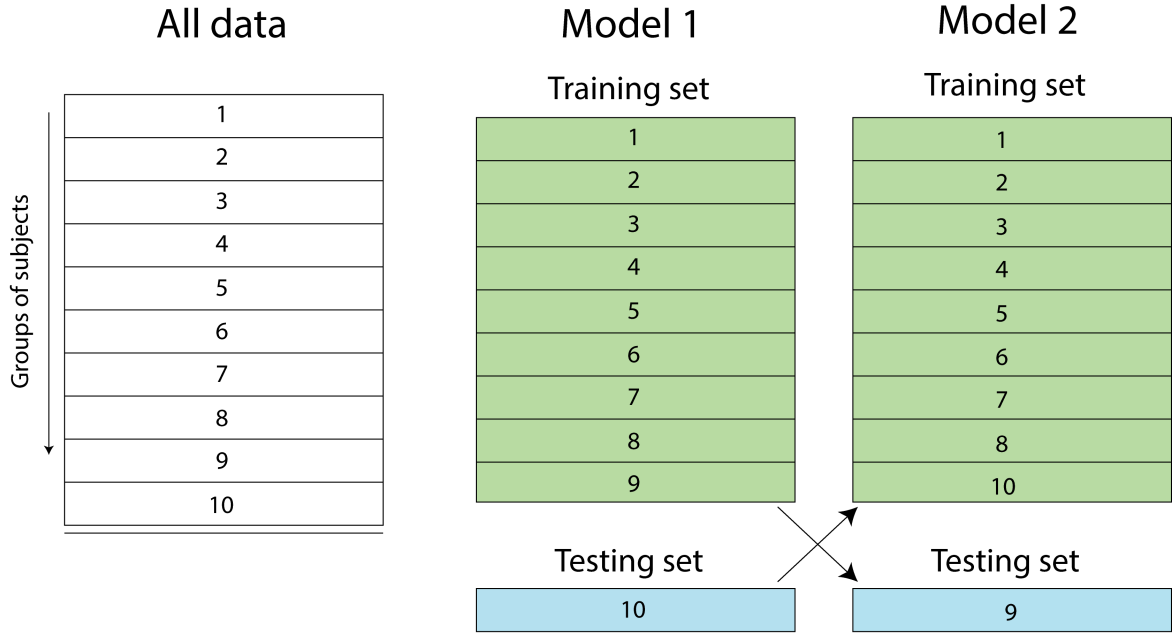
Figure 9.1: 10-fold cross-validation. Each time, a new model is produced, which is trained on a different subset of data. Each group of subjects is fixed, and is used for testing once. The process shown above is repeated 8 more times, so that 10 models arise from cross-validation.

9.2 Machine learning models

In our study, M inputs were used for the machine learning models. This means that the models in our study are operating in an M-dimensional space. However, an M-dimensional space is hard to understand. For this reason, graphical material in this section is shown in a two-dimensional space, which makes it easier to interpret. In this section, three different machine learning models are explained: logistic regression, a bagged random forest and a support vector machine. These models are used after feature selection.

9.2.1 Logistic regression

Logistic regression classification is a method used to separate two groups, based on one or more parameters. When using multiple parameters, this is called a multivariate logistic regression classification. In our study, the probability of a certain subject having LQTS is modeled. The value of this probability ranges between 0 and 1. To this end, the multivariate linear regression formulation (see equation (9.1)) should be transformed, so its minimum and maximum values are 0 and 1. In logistic regression, the logistic function is used to describe the probability of a subject belonging to a certain class (e.g. having LQTS):

$$p_j(x_j) = \frac{1}{1 + e^{-(\alpha + \beta \cdot x_j)}} \tag{9.9}$$

where α is the intercept, β is a row vector containing all coefficients, x_j is a column vector containing all regressors and $p_j(x_j)$ is the probability of this subject having LQTS, see figure 9.2. Consequently, the probability of the same subject being healthy is equal to $1 - p_j(x_j)$. To fit the model, estimates for β_i are sought in such a way that the predicted probability of a subject having LQTS corresponds as closely as possible with the individual's real disease state, also called *maximum likelihood estimation*. In other words: $p_j(x_j)$ should be as close to 0 as possible for all healthy subjects, and as close to 1 as possible for all LQTS patients. This intuition can be specified using the likelihood function:

$$L(\beta) = \prod_{j=1}^{K} \left[\frac{1}{1 + e^{-(\alpha + \beta \cdot x_j)}} \right]^{y_j} \left[1 - \frac{1}{1 + e^{-(\alpha + \beta \cdot x_j)}} \right]^{1-y_j} \tag{9.10}$$

where y_j is the actual class of the subject (healthy or LQTS). Note that if $y_j = 1$, so when the subject has LQTS, the second term is 1. If the subject is healthy, $y_j = 0$, so the first term is 1. So $L(\beta) \leq 1$, and if $L(\beta) = 1$, classification is performed perfectly.
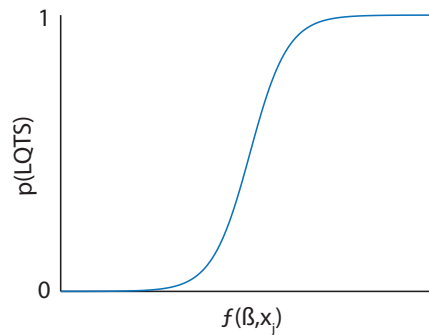

Figure 9.2: The logistic regression function. Probability of having LQTS, plotted against $f(x_j)$.

9.2.2 Support vector machine

In figure 9.3(a), variable x_1 is plotted against x_2. The red dots represent patients with class=1 (e.g. LQTS) and the blue dots represent patients with class=0 (e.g. healthy). To separate both classes, a straight line can be drawn through the data, called a hyperplane. The black lines separate class 0 from class 1, like logistic regression does. However, the question remains which of the three lines separates both classes best when a new sample is introduced. To find a solution to this problem, a maximal margin classifier can be used.
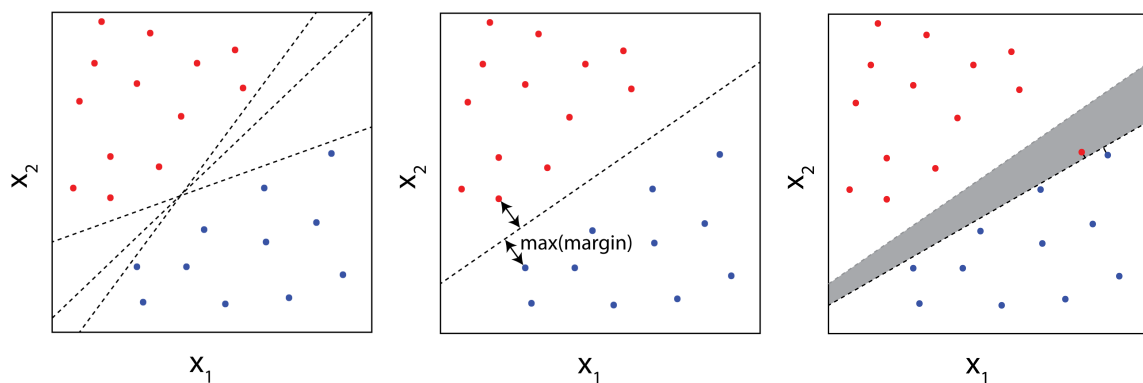
(a) Example of several hyperplanes (dashed lines) to separate classes 0 and 1 as in logistic regression (displayed in blue and red, respectively).

(b) Example of a hyperplane determined with a maximal margin estimator. This hyperplane is the best way to separate both classes.

(c) Example of a hyperplane. One sample has been added to figure 9.3(b), leading to a drastically decreased margin. Consequently, all test subjects in the gray area belonging to class 0 (blue) would be classified wrongly.

Figure 9.3: Examples of hyperplanes.

The maximal margin classifier draws a line through the data which minimizes the distance from the line to its closest points (the *margin*), see figure 9.3(b). By default, a maximal margin classifier would try to separate data perfectly. However, this could cause it to predict new data poorly, see figure 9.3(c). In this figure, one subject is added to figure 9.3(b). This drastically decreases the margin. Consequently, all blue subjects in the gray area would be classified wrongly, in contrast to figure 9.3(b). For this reason, some models not only consider the performance of the system, but also the margin. This kind of model is called a support vector classifier, or a soft margin classifier. If it increases the margin, some of the subjects are allowed to be classified wrongly. The complete working mechanisms of the required regularization techniques are beyond the scope of this thesis. Interestingly, the hyperplane
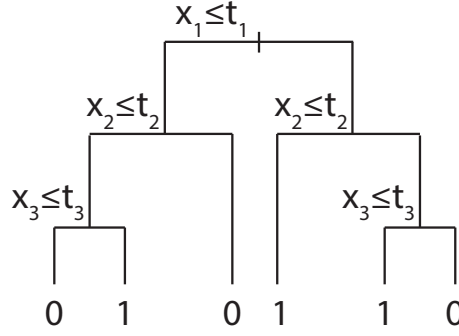
Figure 9.4: Example of a decision tree. The combination of x_1, x_2 and x_3 define how the subject is classified.

following from a support vector classifier depends only upon wrongly classified subjects. Consequently, observations that are far away from the hyperplane do not affect the behaviour of the support vector machine. [24].

9.2.3   Random forest bagged decision tree

Decision tree

Decision trees work through a quite simple-to-grasp means of prediction, resembling human decision making [24]. A number of subsequent binary decisions are made to classify subjects as good as possible. An example of a decision tree is shown in figure 9.4. It consists of a series of splitting rules, starting at the top of the tree. The top split assigns observations with $x_1 \leq t_1$ to the left branch, and observations with $x_1 > t_1$ to the right branch. Subsequently, the value of x_2 is investigated. If $x_2 \leq t_2$, the subject is assigned to the left branch. Otherwise, the subject is assigned to the right branch. Each subsequent binary decision classifies the remaining subjects as well as possible. This is done by iteratively considering all parameters with varying thresholds, and selecting the best parameter to classify subjects in both groups.

In our study, the minimal classification error rate is found before making each decision, which is the recommended criterion when the final classification error rate should be as small as possible [24]. An example of a hyperplane and its corresponding decision tree is shown in figure 9.5. As visible, hyperplanes formed by decision trees always produce rectangular classification areas. The minimal group (*leaf*) size at the bottom of a decision tree should be specified. In our study, the leaf size was set to the default value of 1, since our methods are quite robust for predicting new data, in terms of variance (see sections *Bootstrap-aggregation* and *Random forest*).

Bootstrap-aggregation

Since conventional decision trees tend to predict new data poorly because of their high variance, a more advanced method called bootstrap-aggregation can be used to increase prediction accuracy on test sets. Resulting models are called *bagged decision trees*. Based on the original training set, N training sets ($Z^1, Z^2, Z^3, ... Z^N$) of artificial samples are created, see figure 9.6 [48]. In our study, $N = 500$. Subsequently, N decision trees are trained: one for each artificial dataset. By creating N models, the large variance in conventional decision trees will be divided by N. Hence, bagged decision trees tend to predict new data a lot better than conventional decision trees [24]. To predict the outcome of data in the testing set, the mean of the outcomes of all N models is used.

Random forest

On top of using a bagged decision tree, an additional criterion can be implemented. The N different models of bagged decision trees can be highly correlated with one another, since the strongest predictors tend to be used early in the decision trees. Consequently, averaging all bootstrap-aggregated models will
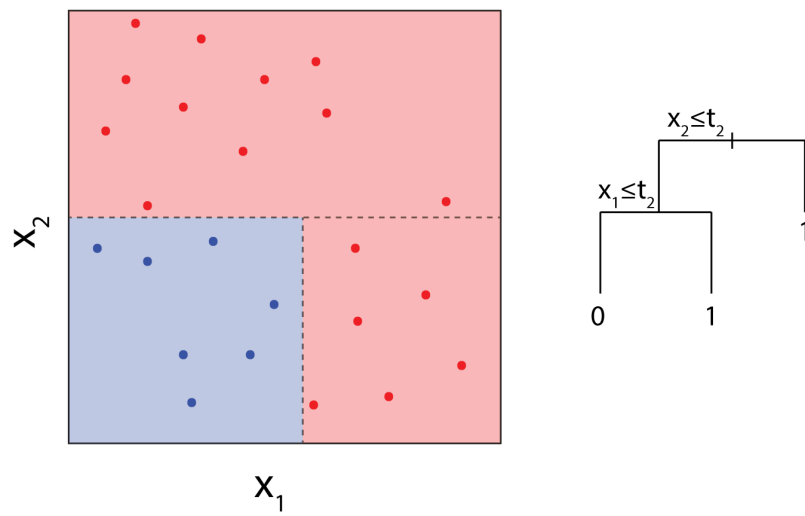
Figure 9.5: Example of a hyperplane and its corresponding decision tree. The combination of x_1 and x_2 define how the subject is classified.
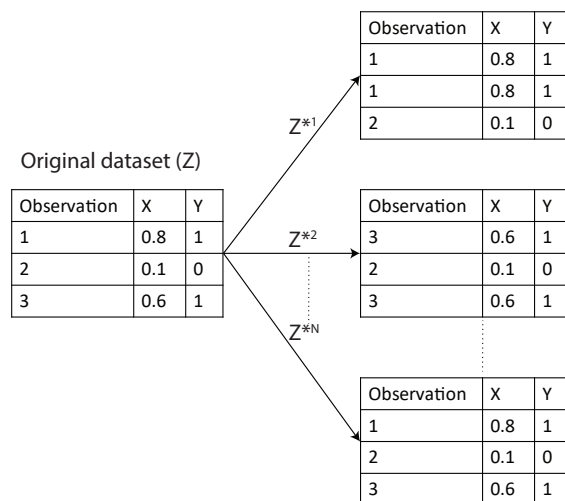
Figure 9.6: Example of bootstrap with replacement. N different artificial training sets are produced, based on samples of the original dataset.

not lead to a significant reduction in bias. [24] To tackle this problem, random forests do not allow the model to consider all predictors at each decision. A random forest only allows each model to consider a subset of predictors (for example, as in our study, $n = \sqrt{p}$, with p being the total number of parameters at each decision). This provides a means of decorrelation. Consequently, the average of the resulting trees will be less variable, increasing prediction accuracy on the testing set. [24]

Bibliography

[1] A. Medeiros-domingo, P. Iturralde-torres, and M.J. Ackerman. Clinical and Genetic Characteristics of Long QT Syndrome. *Revista Española de Cardiología (English Edition)*, 60(7):739–752, 2007.

[2] S. Viskin, P.G. Postema, and Z. Bhuiyan. The Response of the QT Interval to the Brief Tachycardia Provoked by Standing: a bedside test for diagnosing long QT syndrome. *Journal of the American College of Cardiology*, 55(18):1955–1961, 2010.

[3] G. Yan and C. Antzelevitch. Cellular Basis for the Normal T Wave and the Electrocardiographic Manifestations of the Long-QT Syndrome. *Circulation*, 98(18):1928–1936, 1998.

[4] M.P. Andersen, J.Q. Xue, C. Graff, T.B. Hardahl, E. Toft, and J.K. Kanters. A Robust Method for Quantification of IKr-Related T-Wave Morphology Abnormalities Danish National Research Foundation Centre for Cardiac Arrhythmia (DARC), University of Aarhus University Hospital Skejby , Aarhus , Denmark. *Computers in Cardiology*, 34:341–344, 2007.

[5] M.P. Andersen, J.Q. Xue, C. Graff, J.K. Kanters, E. Toft, and J.J. Struijk. New descriptors of T-wave morphology are independent of heart rate. *Journal of Electrocardiology*, 41(6):557–561, 2008.

[6] E. Chorin, O. Havakuk, A. Adler, A. Steinvil, U. Rozovski, C. Van der Werf, et al. Diagnostic value of T-wave morphology changes during " QT stretching " in patients with long QT syndrome. *Heart Rhythm*, 12(11):2263–2271, 2015.

[7] P.J. Schwartz, M. Stramba-badiale, L. Crotti, M. Pedrazzini, A. Besana, G. Bosi, et al. Prevalence of the Congenital Long QT Syndrome. *Circulation*, 120(18):1761–1767, 2009.

[8] D.J. Tester and M.J. Ackerman. Postmortem Long QT Syndrome Genetic Testing for Sudden Unexplained Death in the Young. *Journal of the American College of Cardiology*, 49(2):240–246, 2007.

[9] S.G. Priori, C. Napolitano, P.J. Schwartz, M. Grillo, and J. Nastoli. Association of Long QT Syndrome Loci and Cardiac Events Among Patients Treated With β-Blockers. *American Medical Association*, 292(11):1341, 2004.

[10] S. Viskin, U. Rosovski, A.J. Sands, E. Chen, P.M. Kistler, J.M. Kalman, et al. Inaccurate electrocardiographic interpretation of long QT: The majority of physicians cannot recognize a long QT when they see one. *Heart Rhythm*, 2(6):569–574, 2005.

[11] R.W. Sy, C. van der Werf, I. Chattha, P. Chockalingam, A. Adler, J. Healey, et al. Derivation and validation of a simple exercise-based algorithm for prediction of genetic testing in relatives of LQTS probands. *Circulation*, 124:2187–2194, 2011.

[12] S. Baskar and P.F. Aziz. Genotype-phenotype correlation in long QT syndrome. *Global cardiology science & practice*, 2015(2):26, 2015.

[13] G.M. Vincent. Role of DNA testing for diagnosis, management, and genetic screening in long QT syndrome, hypertrophic cardiomyopathy, and Marfan syndrome. *Heart*, 86:12–14, 2001.

[14] S.G. Priori, C. Napolitano, and P.J. Schwartz. Low Penetrance in the Long-QT Syndrome. *Circulation*, 99:529–533, 1999.

[15] D. Ricciardi, I. Cavallari, A. Creta, G. Di Giovanni, V. Calabrese, N. Di Belardino, , et al. Impact of the high-frequency cutoff of bandpass filtering on ECG quality and clinical interpretation: A comparison between 40 Hz and 150 Hz cutoff in a surgical preoperative adult outpatient population. *Journal of Electrocardiology*, 49(5):691–695, 2016.

[16] Pan, J. and Tompkins, Willis. J. A real-time qrs detection algorithm. *IEEE transactions on biomedical engineering*, 3:230–236, 1985.

[17] C. Orphanidou, T. Bonnici, P. Charlton, D. Clifton, D. Vallance, and L. Tarassenko. Signal Quality Indices for the Electrocardiogram and Photoplethysmogram: Derivation and Applications to Wireless Monitoring. *IEEE journal of biomedical and health informatics*, 2194(c):832–838, 2014.

[18] B.J.M. Hermans, A.S. Vink, F.C. Bennis, L.H. Filippini, V.M.F. Meijborg, A.A.M. Wilde, et al. The development and validation of an easy to use automatic QT-interval algorithm. *PLoS ONE*, 12(9):1–14, 2017.

[19] A. Mincholé, R. Ariga, S. Neubauer, H. Watkins, and B. Rodr. Electrocardiographic abnormalities of hypertrophic cardiomyopathy. *Computing in cardiology*, 41:397–400, 2014.

[20] R.L. Verrier and H. Huikuri. Tracking interlead heterogeneity of R- and T-wave morphology to disclose latent risk for sudden cardiac death. *Heart Rhythm*, 14(10):1466–1475, 2017.

[21] A. Oehler, T. Feldman, C. Hendrikson, and L.G. Tereshchenko. QRS-T Angle: A Review. *Annals of noninvasive electrocardiography*, 19(6):534–542, 2014.

[22] A.Y. Tan, Br.D. Nearing, M. Rosenberg, R. Nezafat, M.E. Josephson, and R.L. Verrier. Interlead heterogeneity of R- and T-wave morphology in standard 12-lead ECGs predicts sustained ventricular tachycardiafibrillation and arrhythmic death in patients with cardiomyopathy. *Journal of Cardiovascular Electrophysiology*, 28(11):1324–1333, 2017.

[23] J. Friedman, T. Hastie, and R. Tibshirani. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal Of Statistical Software*, 35(1):1–15, 2010.

[24] G. James, D. Witten, T. Hastie, and R. Tibshirani. Linear Model Selection and Regularization. In *An Introduction to Statistical Learning: with Applications in R (ISLR)*, pages 205–268. 2013.

[25] R.M.A. Ter Bekke, K.H. Haugaa, A. Van Den Wijngaard, J.M. Bos, M.J. Ackerman, T. Edvardsen, et al. Electromechanical window negativity in genotyped long-QT syndrome patients: Relation to arrhythmia risk. *European Heart Journal*, 36(3):179–186, 2014.

[26] A. Khositseth, J. Hejlik, W.K. Shen, and M.J. Ackerman. Epinephrine-induced T-wave notching in congenital long QT syndrome. *Heart Rhythm*, 2(2):141–146, 2005.

[27] W. Zareba, A.J. Moss, S. le Cessie, and W. J. Hall. T wave alternans in idiopathic long QT syndrome. *Journal of the American College of Cardiology*, 23(7):1541–1546, 1994.

[28] S.B. Platt, J.M. Vijgen, P. Albrecht, G.F. Van Hare, M.D. Carlson, and D.S. Rosenbaum. Occult T wave alternans in long QT syndrome. *Journal of Cardiovascular Electrophysiology*, 7(2):144–148, 1996.

[29] J.N. Johnson and M.J. Ackerman. QTc: how long is too long? *British Journal of Sports Medicine*, 43(9):657–662, 2009.

[30] S. Wong, G. Kervio, M. Altuve, F. Carre, and G. Carrault. Comparing six QT correction methods in an athlete population. *Computing in Cardiology*, 39:585–588, 2012.

[31] S. Luo, K. Michler, P. Johnston, and P.W. MacFarlane. A comparison of commonly used QT correction formulae: The effect of heart rate on the QTc of normal ECGs. *Journal of Electrocardiology*, 37:81–90, 2004.

[32] Y. Lokhandwala and S. C. Toal. The fallacies of QT correction. *Indian Pacing and Electrophysiology Journal*, 3(4):185–186, 2003.

[33] P.J. Schwartz and L. Crotti. QTc behavior during exercise and genetic testing for the long-qt syndrome. *Circulation*, 124(20):2181–2184, 2011.

[34] A.C. Guyton and J.E. Hall. The normal electrocardiogram. In William Schmitt and Rebecca Gruliow, editors, *Textbook of Medical Physiology*, chapter 11, pages 123–130. Philadelphia, 11 edition, 2006.

[35] E.L. Boulpaep and W.F. Boron. Cardiac Electrophysiology and the Electrocardiogram. In *Medical Physiology*, chapter 21, pages 953–994. Elsevier Inc., Philadelphia, 2 edition, 2012.

[36] M. Liu, K.-C. Yang, and S.C. Dudley. Chapter sixteen - cardiac sodium channel mutations: Why so many phenotypes? In Robert J. French and Sergei Yu. Noskov, editors, *Na Channels from Phyla to Function*, volume 78 of *Current Topics in Membranes*, pages 513 – 559. Academic Press, 2016.

[37] C. VanPutte, J. Regan, and A. Russo. Cardiovascular System: The Heart. In Mandy Clark and Deb Debord, editors, *Seeley's Anatomy & Physiology*, chapter 20, pages 665–709. McGraw-Hill, New York, 10 edition, 2014.

[38] R.K. Hobbie and B.J. Roth. The Exterior Potential and the Electrocardiogram. In *Intermediate Physics for Medicine and Biology*, chapter 7, pages 177–201. Springer Science+Business Media, LLC, New York, fourth edi edition, 2007.

[39] R.A. Rhoades and G.A. Tanner. The Electrical Activity of the Heart. In *Medical Physiology*, chapter 13, pages 219–236. Lippincott Williams & Wilkins, Philadelphia, 2 edition, 2003.

[40] E. Frank. An accurate, clinically practical system for spatial vectorcardiography. *Circulation*, 13:737–749, 1956.

[41] J.N. Weiss, A. Garfinkel, and H.S. Karagueuzian. Early afterdepolarizations and cardiac arrhythmias. *Heart Rhythm*, 7(12):1891–1899, 2011.

[42] J. Edhouse and F. Morris. Broad complex tachycardia - Part II. In *ABC of Clinical Electrocardiography*, chapter 7, pages 25–28. BMJ Books, London, 1 edition, 2003.

[43] P.J. Schwartz. Idiopathic long QT syndrome: Progress and questions. *American Heart Journal*, 109(2):399–411, 1985.

[44] A.J. Moss, W. Zareba, J. Benhorin, E.H. Locati, and K.W. Timothy. ECG T-wave Patterns in Genetically Distinct Forms of the Herieditary Long QT Syndrome.pdf. *Circulation*, 92(10):2929–2934, 1995.

[45] E. Dausse, M. Berthet, I. Denjoy, X. André-Fouet, C. Cruaud, M. Bennaceur, et al. A mutation in HERG associated with notched T waves in long QT syndrome. *Journal of Molecular and Cellular Cardiology*, 28(8):1609–1615, 1996.

[46] C. Antzelevitch. Drug-Induced Channelopathies. In *Cardiac Electrophysiology: from cell to bedside*, chapter 19, pages 195–204. Elsevier, Philadelphia, 5 edition, 2009.

[47] H. Zou and T. Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 67(2):301–320, 2005.

[48] L. Breiman. Random forests. In Robert E. Schapire, editor, *Machine Learning*, volume 45, pages 5–32. Kluwer Academic Publishers, Berkeley, California, 2001.

Part III

Appendices

Chapter A

Supplemental info on methods

A.1 Linear differentiator

A simple linear differentiator was used to calculate the derivative of several signals in our study. The differentiator works as follows:

$$\dot{f}(t) = \frac{f(t-25) - f(t+25)}{(t-25) - (t+25)} \tag{A.1}$$

where t is the time in ms and $f(t)$ is the signal to be differentiated.

A.2 Skewness

After T-wave transformation, the skewness was calculated as follows:

$$\text{Skewness} = \frac{E(x-\mu)^3}{\sigma^3} \tag{A.2}$$

where x is the T-wave, μ is the mean of the signal, σ is the standard deviation of x, and $E(t)$ represents the expected value of the quantity t. The skewness can also be written as:

$$\text{Skewness} = \frac{\frac{1}{n}\sum_{i=0}^{n-1}(x_i - \bar{x})^3}{\left(\frac{1}{n}\sum_{i=0}^{n-1}(x_i - \bar{x})^2\right)^{\frac{3}{2}}} \tag{A.3}$$

where n is the number of samples and $\bar{x}$ the mean of value of all samples. The skewness of a distribution depends on the ratio of its mean to its median. When mean < median, the skewness is smaller than 0, leading to a left-tailed distribution. When mean = median, this leads to a skewness of 0 and a symmetrical distribution. When median < mean, the skewness is larger than 0, which makes the distribution right-tailed.

A.3 Kurtosis

The kurtosis, after transforming the T-wave, was calculated as follows:

$$\text{Kurtosis} = \frac{E(x-\mu)^4}{\sigma^4} \tag{A.4}$$

where μ is the mean of ECG, σ is the standard deviation of x, and $E(t)$ represents the expected value of the quantity t. The kurtosis can also be written as:

$$\text{Kurtosis} = \frac{\frac{1}{n}\sum_{i=0}^{n-1}(x_i - \bar{x})^4}{\left(\frac{1}{n}\sum_{i=0}^{n-1}(x_i - \bar{x})^2\right)^2} \tag{A.5}$$

where n is the number of samples and $\bar{x}$ the mean of value of all samples. A higher kurtosis means that the distribution is more 'peaked', and a lower kurtosis means a less peaked distribution. An average kurtosis of 3 belongs to a normal distribution.

A.4 Spatial peak QRS-T angle

To calculate the spatial peak QRS-T angle, the maximal QRS complex magnitude (QRS_p) and T peak magnitude (T_p) were found in the VCG. The spatial peak QRS-T angle was calculated using the definition of the normalized inner product of spatial peak QRS and T vectors, according to the following equation:

$$\text{SP QRS-T angle} = arccos \left(\frac{\overline{QRS_p} \cdot \overline{T_p}}{|\overline{QRS_p}| \cdot |\overline{T_p}|} \right) \tag{A.6}$$

where $\overline{QRS_p}$ is the vector at the maximal magnitude of the QRS complex, and $\overline{T_p}$ the vector at the maximal magnitude of the T-wave.

A.5 Spatial mean QRS-T angle

To calculate the spatial mean QRS-T angle of the average complex, the mean QRS complex vector ($\overline{QRS_m}$) and mean T peak vector ($\overline{T_m}$) were calculated from the VCG, as follows [21]:

$$\overline{QRS_m} = \frac{\sum\limits_{t=Q_{AC}}^{R_{AC}+50} \overline{x_t}}{R_{AC} + 50 - Q_{AC}} \tag{A.7}$$

$$\tag{A.8}$$

$$\overline{T_m} = \frac{\sum\limits_{t=T_{start}}^{T_{end}} \overline{x_t}}{T_{end} - T_{start}} \tag{A.9}$$

where $\overline{x_t}$ is the 3-dimensional VCG signal at each time, R_{AC} is the R-peak, Q_{AC} is the Q-wave, T_{end} is the end of the T wave, T_{start} is the start of the T wave (see section 2.4), $\overline{QRS_m}$ is the mean vector during the QRS-complex and $\overline{T_m}$ is the mean vector during the T-wave. Subsequently, the spatial mean QRS-T angle was calculated using the definition of the normalized inner product of spatial mean QRS and T vectors, according to the following equation [21]:

$$\text{SM QRS-T angle} = arccos \left(\frac{\overline{QRS_m} \cdot \overline{T_m}}{|\overline{QRS_m}| \cdot |\overline{T_m}|} \right) \tag{A.10}$$

Chapter B

Features per lead

Table B.1: Table indicating which leads are necessary to calculate features. TH: T-wave heterogeneity. SP: Spatial peak. SM: Spatial mean. RH: R wave heterogeneity. TpTe: T-wave peak to T-wave end interval. ^{T}: T-wave morphology feature. *: Global feature. †: Feature based on multiple leads: only one value results from all denoted leads.

Feature	VR	VL	VF	V1	V2	V3	V4	V5	V6	RMS	VCG X	VCG Y	VCG Z	VCG magnitude	PCA1
Area^T	X	X	X	X	X	X	X	X	X						
Absolute area^T	X	X	X	X	X	X	X	X	X						
Biphasicness^T	X	X	X	X	X	X	X	X	X						
Amplitude^T	X	X	X	X	X	X	X	X	X						
Skewness^T	X	X	X	X	X	X	X	X	X						
Kurtosis^T	X	X	X	X	X	X	X	X	X						
Notch score^T	X	X	X	X	X	X	X	X	X						
Asymmetry score^T	X	X	X	X	X	X	X	X	X						
QRS amplitude	X	X	X	X	X	X	X	X	X						
TH(V1-V3)^{T†}				X	X	X									
TH(V4-V6)^{T†}							X	X	X						
SP QRS-T angle*†											X	X	X	X	
SM QRS-T angle*†											X	X	X	X	
RH(V1-V3)†				X	X	X									
RH(V4-V6)†							X	X	X						
TpTe*															X
R to T interval*										X					X
Length*															X

Chapter C

Log-likelihood function

The following equation shows how the logarithm of the maximum likelihood function (equation I.(9.7)) leads to a logistic regression formulation combined with elastic net (equation I.(9.3)).
For simplicity, $S(\alpha, \beta, x_j) = \frac{1}{1+e^{-(\alpha+\beta \cdot x_j)}}$.

$$\ln \left(\prod_{j=1}^{K} \left[S(\alpha, \beta, x_j) \right]^{y_j} \cdot \left[1 - S(\alpha, \beta, x_j) \right]^{1-y_j} \right)$$

$$= \sum_{j=1}^{K} \left[y_j \cdot \ln \left(S(\alpha, \beta, x_j) \right) + (1 - y_j) \cdot \ln \left(1 - S(\alpha, \beta, x_j) \right) \right]$$

$$= \sum_{j=1}^{K} \left[y_j \cdot \ln \left(\frac{1}{1 + e^{-(\alpha+\beta \cdot x_j)}} \right) + (1 - y_j) \cdot \ln \left(1 - \frac{1}{1 + e^{-(\alpha+\beta \cdot x_j)}} \right) \right]$$

$$= \sum_{j=1}^{K} \left[y_j \cdot \ln \left(\frac{1}{1 + e^{-(\alpha+\beta \cdot x_j)}} \right) + (1 - y_j) \cdot \ln \left(1 - \frac{e^{-(\alpha+\beta \cdot x_j)}}{1 + e^{-(\alpha+\beta \cdot x_j)}} \right) \right]$$

$$= \sum_{j=1}^{K} \left[\ln \left(1 - \frac{e^{-(\alpha+\beta \cdot x_j)}}{1 + e^{-(\alpha+\beta \cdot x_j)}} \right) + y_j \left(\ln \left(\frac{1}{1 + e^{-(\alpha+\beta \cdot x_j)}} \right) - \ln \left(\frac{e^{-(\alpha+\beta \cdot x_j)}}{1 + e^{-(\alpha+\beta \cdot x_j)}} \right) \right) \right]$$

$$= \sum_{j=1}^{K} \left[\ln \left(1 - \frac{e^{-(\alpha+\beta \cdot x_j)}}{1 + e^{-(\alpha+\beta \cdot x_j)}} \cdot \frac{e^{(\alpha+\beta \cdot x_j)}}{e^{(\alpha+\beta \cdot x_j)}} \right) + y_j \left(\ln \left(\frac{1}{1 + e^{-(\alpha+\beta \cdot x_j)}} \cdot \frac{1 + e^{-(\alpha+\beta \cdot x_j)}}{e^{-(\alpha+\beta \cdot x_j)}} \right) \right) \right]$$

$$= \sum_{j=1}^{K} \left[\ln \left(\frac{1}{1 + e^{(\alpha+\beta \cdot x_j)}} \right) + y_j \left(\ln \left(\frac{1}{e^{-(\alpha+\beta \cdot x_j)}} \right) \right) \right]$$

$$= \sum_{j=1}^{K} \left[\ln(1) - \ln \left(1 + e^{(\alpha+\beta \cdot x_j)} \right) + y_j(\ln(1) - \ln(e^{-(\alpha+\beta \cdot x_j)})) \right]$$

$$= \sum_{j=1}^{K} \left[- \ln \left(1 + e^{-(\alpha+\beta \cdot x_j)} \right) + \beta \cdot y_j \cdot x_j \right].$$

Chapter D

Included parameters in final model

Table D.1: Included features and β-values in final model. These values are retrieved after standardization of all parameters.

Feature	β-value
Age	-0.012
QTc Bazett	0.0063
QTc Fridericia	0.0037
QTc Framingham	0.0037
QTc Hodges	0.0043
Area in V2	1.8E-6
Area in V3	2.0E-6
Absolute area in V1	3.9E-6
Absolute area in V3	3.4E-6
Biphasicness in VR	-0.15
Biphasicness in VF	-6.2E-4
Amplitude in V6	-1.2E-4
Length	0.014
Time to onset	0.0091
Skewness in VR	-0.73
Skewness in VL	-0.17
Skewness in VF	-0.55
Skewness in V5	-.48
Skewness in V6	-0.082
Kurtosis in VL	0.29
Kurtosis in V3	0.26
Asymmetry in VR	1.01
Asymmetry in VF	0.0075
Asymmetry in V3	0.24
Asymmetry in V4	0.043
T_{pk} to T_{end} interval	0.48
Spatial peak QRS-T angle	0.0021
R-peak heterogeneity in V4-V6	-8.3E-5
T-wave heterogeneity in V4-V6	6.89E-4