

RESEARCH ARTICLE

Investigating automated regression models for estimating left ventricular ejection fraction levels in heart failure patients using circadian ECG features

Sona M. Al Younis^{1*}, Leontios J. Hadjileontiadis^{1,2}, Aamna M. Al Shehhi¹, Cesare Stefanini³, Mohanad Alkhodari^{1,4}, Stergios Soulaïdopoulos⁵, Petros Arsenos⁵, Ioannis Doundoulakis⁵, Konstantinos A. Gatzoulis⁵, Konstantinos Tsioufis⁵, Ahsan H. Khandoker¹

1 Department of Biomedical Engineering, Healthcare Engineering Innovation Centre (HEIC), Khalifa University, Abu Dhabi, United Arab Emirates, **2** Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece, **3** Creative Engineering Design Lab at the BioRobotics Institute, Applied Experimental Sciences Scuola Superiore Sant'Anna, Pontedera (Pisa), Italy, **4** Cardiovascular Clinical Research Facility, Radcliffe Department of Medicine, University of Oxford, Oxford, United Kingdom, **5** First Cardiology Department, School of Medicine, "Hippokration" General Hospital, National and Kapodistrian University of Athens, Athens, Greece

* 100058882@ku.ac.ae



OPEN ACCESS

Citation: Al Younis SM, Hadjileontiadis LJ, Al Shehhi AM, Stefanini C, Alkhodari M, Soulaïdopoulos S, et al. (2023) Investigating automated regression models for estimating left ventricular ejection fraction levels in heart failure patients using circadian ECG features. *PLoS ONE* 18(12): e0295653. <https://doi.org/10.1371/journal.pone.0295653>

Editor: Amirmohammad Khalaji, Tehran University of Medical Sciences, ISLAMIC REPUBLIC OF IRAN

Received: September 11, 2023

Accepted: November 22, 2023

Published: December 11, 2023

Copyright: © 2023 Al Younis et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data set for the American patient cohort was obtained from the archives of the Intercity Digital Electrocardiography (ECG) Alliance (IDEAL) study of the University of Rochester Medical Center Telemetric and Holter ECG Warehouse (THEW). That is accessible as a publicly published dataset through the link: <http://www.thew-project.org/databases.htm>, which is accessed by registering first through the link: http://www.thew-project.org/register_main.htm

Abstract

Heart Failure (HF) significantly impacts approximately 26 million people worldwide, causing disruptions in the normal functioning of their hearts. The estimation of left ventricular ejection fraction (LVEF) plays a crucial role in the diagnosis, risk stratification, treatment selection, and monitoring of heart failure. However, achieving a definitive assessment is challenging, necessitating the use of echocardiography. Electrocardiogram (ECG) is a relatively simple, quick to obtain, provides continuous monitoring of patient's cardiac rhythm, and cost-effective procedure compared to echocardiography. In this study, we compare several regression models (support vector machine (SVM), extreme gradient boosting (XGBOOST), gaussian process regression (GPR) and decision tree) for the estimation of LVEF for three groups of HF patients at hourly intervals using 24-hour ECG recordings. Data from 303 HF patients with preserved, mid-range, or reduced LVEF were obtained from a multicentre cohort (American and Greek). ECG extracted features were used to train the different regression models in one-hour intervals. To enhance the best possible LVEF level estimations, hyper-parameters tuning in nested loop approach was implemented (the outer loop divides the data into training and testing sets, while the inner loop further divides the training set into smaller sets for cross-validation). LVEF levels were best estimated using rational quadratic GPR and fine decision tree regression models with an average root mean square error (RMSE) of 3.83% and 3.42%, and correlation coefficients of 0.92 ($p < 0.01$) and 0.91 ($p < 0.01$), respectively. Furthermore, according to the experimental findings, the time periods of midnight-1 am, 8–9 am, and 10–11 pm demonstrated to be the lowest RMSE values between the actual and predicted LVEF levels. The findings could potentially lead to the

The Greek patient cohort was obtained from the PRESERVE EF study with patients enrolled across seven cardiology departments in Greece, this data set is available upon request through: doi: [10.1093/eurheartj/ehz260](https://doi.org/10.1093/eurheartj/ehz260). Machine learning networks, final trained models, and regression codes are available at: https://github.com/sonametep2/Sona_LVEFRegression_2023.

Funding: Research reported in this publication was supported by grant (award number: 8474000132) from the Healthcare Engineering Innovation Center (HEIC) at Khalifa University, Abu Dhabi, UAE. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

development of an automated screening system for patients with coronary artery disease (CAD) by using the best measurement timings during their circadian cycles.

Introduction

Heart failure (HF) affects approximately 26 million people worldwide [1], which is an increasing global burden on cardiologists with 3.5 million new patients yearly [2]. At 55 years of age, the lifetime risks of heart failure for men and women are 29% and 33%, respectively [3]. According to estimates, 480,000 adults (aged 18 or older) suffer from cardiac failure, which accounts for 2.1% of the population in adults [4]. While there are slight variations in the definitions of heart failure used in the current practice guidelines from the American College of Cardiology (ACC)/American Heart Association (AHA) [5], Heart Failure Association (HFA)/European Society of Cardiology (ESC) [6], and Japanese Heart Failure Society (JHFS) [7], the overall concepts and criteria for HF classification into different stages are based on the severity and progression of the condition and the identifiable signs and symptoms including edema/fluid retention, dyspnea, activity intolerance, and fatigue. They also emphasize the presence of structural or functional heart disease as a prerequisite for the diagnosis. Coronary artery disease (CAD) is the leading cause of HF, followed by hypertension, diabetes, valvular heart disease, and cardiomyopathy [8]. Left Ventricle Ejection Fraction (LVEF) levels, which is a hemodynamic term for the fraction of ventricular volume ejected per heartbeat [9], was further highlighted as important indicators for diagnosis, prognosis, and treatment of HF patients [10, 11], physical examination, patient history, and clinical tests [12].

According to the American Society of Echocardiography and the European Association of Cardiovascular Imaging (ASE/EACVI) [13, 14], systolic dysfunction, usually called heart failure with reduced ejection fraction (HFrEF), is the clinical manifestation of HF symptoms with a resulting LVEF of less than 50%. If the measured EF is more than 55% it is a diastolic dysfunction, usually called heart failure with preserved ejection fraction (HFpEF). Nevertheless, if LVEF is slightly reduced (LVEF 50–55%), the failure category is heart failure with mid-range ejection fraction (HFmEF). Due to the HF etiology, the HFmEF category's narrower range is seen as a changeable criterion for this category. Different cut-off values for the classification of HF are advised by other standards, such as the ESC [6] and JHFS [7]; the cut-off for HFrEF is as low as 40%. According to the literature, there are no rigid guidelines, and the treatment is only tangentially related to LVEF and clinical presentation. However, based on the ESC criteria, patients in the mid-range group between 40 and 49% showed that 90% of patients either got better or got worse, with only 10% of instances remaining unaltered.

LVEF is estimated from nuclear medicine scans [15], computerized tomography (CT) [16], and cardiac catheterization [17]. However, cardiovascular magnetic resonance (CMR) [18] imaging and echocardiography [19] are the most reliable and widely used methods to evaluate LVEF. Despite the use of non-ionizing radiations and high estimation accuracy, potential limitations are associated with these techniques, i.e., cost and accessibility, operator dependency, patient contraindications, time and patient cooperation, and image interpretation challenges. Electrocardiography (ECG) presents a supplementary tool that is both accessible and cost-effective for the evaluation of LVEF. The choice of ECG as a method for LVEF assessment is substantiated by its capability to capture key physiological connections reflective of cardiac performance. ECG features, such as QRS duration, ST segment changes, and T-wave abnormalities, are indicative of electrical conduction abnormalities, myocardial ischemia, and ventricular repolarization anomalies. These electrical manifestations are intricately linked to the

mechanical processes governing ventricular contraction and relaxation. Notably, alterations in LVEF can influence cardiac electrical activities, giving rise to deviations in ECG patterns. Consequently, by comprehensively analysing ECG signals and their derived parameters, it becomes plausible to discern the interplay between these electrical indicators and LVEF, offering valuable insights into cardiac function without the necessity of more invasive or costly procedures.

ECG is one of the most widely used non-invasive diagnostic tools for tracking the physiological activities of the heart throughout time. Many cardiovascular disorders, including atrial fibrillation, myocardial infarction, premature contractions of the ventricles or atria, and congestive heart failure, can be diagnosed with the help of ECG data [20–25]. Consequently, automatic and accurate ECG data analysis has become a popular study area, especially using artificial intelligence (AI) tools. The diagnostic golden rules have been used for automatic ECG analysis. This process consists of two steps that involve human specialists to create meaningfully featuring from raw ECG data; these features could be categorized into coefficients of variation and density histograms, statistical features (such as heart rate variability), frequency-domain features, time-domain features, and sample entropy. Based on ECG and machine learning tools, several researchers classified patients exhibiting congestive HF from normal [26, 27]. In contrast, deep learning with one lead ECG signal was used to categorize HF patients into CAD, Myocardial Infarction (MI), and congestive HF classes [20].

However, there is still a limited understanding of the intricate relationship between circadian ECG wave characteristics and the estimation of LVEF in patients with heart failure. Moreover, it would be valuable to present a more accessible alternative approach for evaluating LVEF that doesn't necessitate extensive expertise or costly equipment. Along the same lines, machine learning techniques, including deep learning, can play a crucial role in comprehending the intricate ECG features present in-patient records, ultimately leading to improved assessment of HF. Hence, the objectives of this study were:

1. To explore the potential of machine learning regression models trained on features extracted from the ECG waveform.
2. In line with the ASE/EACVI guidelines, to accurately estimate LVEF levels in HFpEF, HFmEF, and HFrfEF HF categories.
3. To analyse the cardiovascular system dynamics in HF patients hour by hour throughout a 24-hour day.
4. To emphasize the heart's 24-hour circadian functionality (correlated with ECG features) to recommend the best times for a correct estimation of LVEF levels, offering a suitable screening strategy with the best window of possibilities for halting heart failure progression.

Materials and methods

Dataset and patients' enrollment

The complete procedure followed in this study is illustrated in Fig 1. This study used two datasets that comprised clinical data from patient groups of Greeks and Americans. Patients with HF, specifically CAD, and ages ranging from 33 to 88 years old ($n = 303$) were included in both datasets. According to the ASE/EACVI recommendations, these patients were split into 129 HFpEF, 92 HFmEF, and 82 HFrfEF groups.

Patients from seven cardiology departments in Greece who were engaged in the PRESERVE EF trial contributed to the Greek patient cohort [28]. The enrolment protocol of patients for the PRESERVE EF study (clinicaltrials.gov/identifier NCT02124018) was approved by the ethics

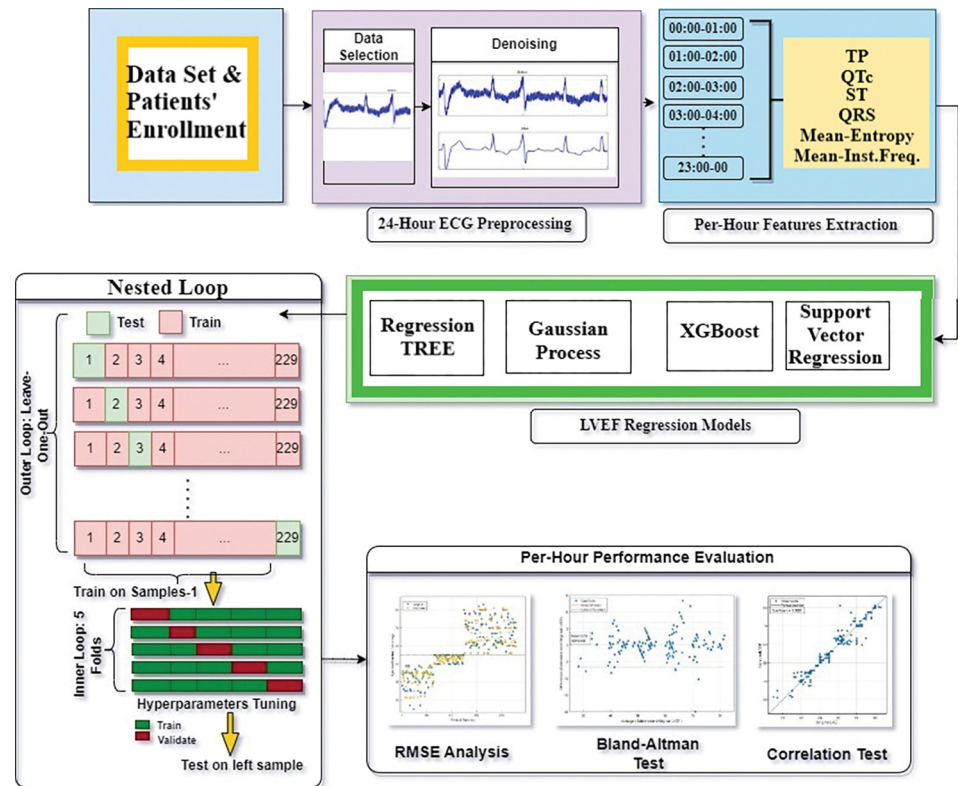


Fig 1. An illustration of the overall procedure followed in this study.

<https://doi.org/10.1371/journal.pone.0295653.g001>

committee at each of the seven selected cardiology departments at Greece and was endorsed by the Hellenic Society of Cardiology. The Hellenic Society of Cardiology established and maintained a database [29]. Each patient signed a consent form before enrolling in the trial at each cardiology department. Patients needed to meet the following requirements to be eligible for enrolment: (1) having a post-angiographically proven MI at least 40 days after the event or 90 days after any CABG surgeries, if applicable; (2) being revascularized; (3) not being revascularized but lacking evidence of any active ischemia in the previous six months; and (4) completing an effective course of medical treatment.

Additionally, any patient who had a secondary prevention indication for the placement of an implantable cardioverter defibrillator (ICD), a permanent pacemaker, persistent, long-standing persistent, or permanent atrial fibrillation, any neurological symptoms of syncope or pre-syncope within the previous six months, or the presence of any systemic illnesses like liver failure, renal diseases, rheumatic diseases, thyroid dysfunction, or cancer was excluded from the study. For all patients, the American Society of Echocardiography's recommendations are followed when doing echocardiographic tests, and a GE Healthcare GETEMED CardioDay Holter system (recorder CardioMem CM4000 and software CardioDay v. 2.4, GE Healthcare, Fairfield, CT, USA) is used.

The Intercity Digital Electrocardiography (ECG) Alliance (IDEAL) study archives at the University of Rochester Medical Centre Telemetric and Holter ECG Warehouse (THEW) were used to identify the American patient cohort [30]. The database enrolment protocol complied with the Declaration of Helsinki and Title 45, U.S. Code of Federal Regulations, Part 46, Protection of Human Subjects (revised: November 13, 2001-effective: December 13, 2001).

Table 1. Clinical characteristics of the heart failure patients based on their LVEF categories.

| Clinical Variables | Overall Subjects | LVEF Categories | | | p-value |
|--|------------------|-----------------|---------------|---------------|---------|
| | | HFpEF (n =) | HFmEF (n =) | HFrfEF (n =) | |
| Patients (n) | 229 | 105 | 60 | 64 | - |
| LVEF (%) | 25–82 | 56–82 | 50–55 | 25–49 | <0.001 |
| (Mean±Std) | (56.16±11.79) | (66.63±6.7) | (52.82±2.24) | (42.12±5.52) | - |
| Gender (M/F) | 196/33 | 90/15 | 49/11 | 57/7 | 0.237 |
| Age (yrs.) | 35–88 | 35–79 | 40–88 | 40–80 | 0.110 |
| (Mean±Std) | (58.29±10.75) | (57.14±10.4) | (60.75±10.88) | (57.86±10.97) | - |
| BMI (kg/m ²) | 17.99–37.89 | 19.72–36.33 | 17.99–37.65 | 20.8–37.89 | 0.133 |
| (Mean±Std) | (27.15±3.62) | (26.66±3.44) | (27.49±3.89) | (27.61±3.59) | - |
| Smoking (Yes/No) | 167/62 | 72/33 | 48/12 | 47/17 | 0.757 |
| Diabetes (Yes/No) | 29/200 | 7/98 | 7/53 | 15/49 | <0.001 |
| Hypertension (Yes/No) | 113/116 | 51/54 | 31/29 | 31/33 | 0.587 |
| Angina (Yes/No) | 159/70 | 83/22 | 37/23 | 39/25 | 0.189 |
| VT (Yes/No) | 19/210 | 9/96 | 2/58 | 8/56 | 0.088 |
| Prior-MI (Yes/No) | 150/79 | 54/51 | 44/16 | 52/12 | <0.001 |
| Medications (Beta-Blockers, ACE Inhibitor, Anti-Arrhythmic, Diuretic) | | | | | |
| Beta-Blockers (Yes/No) | 179/50 | 80/25 | 46/14 | 53/11 | 0.478 |
| ACE Inhibitor (Yes/No) | 68/161 | 30/75 | 15/45 | 23/41 | 0.611 |
| Anti-Arrhythmic (Yes/No) | 7/222 | 3/102 | 4/56 | 0/64 | 0.166 |
| Diuretic (Yes/No) | 42/187 | 1/104 | 19/41 | 22/42 | <0.001 |

BMI: Body Mass Index, VT: Ventricle Tachycardia, MI: Myocardial Infarction, ACE: Angiotensin converting enzyme.

<https://doi.org/10.1371/journal.pone.0295653.t001>

Additionally, the IDEAL protocol was authorised by the University of Rochester's research subject review board [31]. Before taking part in the trial, each patient gave their written consent.

The following requirements had to be met to qualify for enrolment in the IDEAL study: (1) having evidence of either a prior MI or exercise-induced ischemia; (2) being in the stable phase of ischemic heart disease at least two months after the last event; (3) not having been given a congenital heart failure diagnosis; and (4) having sinus rhythm. Additionally, any patients with dilated cardiomyopathy—defined as having a left ventricular diameter (LVD) > 60 mm and an ejection fraction (EF) 40%—congenital heart failure (CHF), coronary artery bypass grafting (CABG) surgery, non-sinus rhythm, and any cerebral, severely hepatic, or malignant diseases—were disqualified from the trial. All patients underwent an echocardiography examination to determine their LVEF levels. In addition, a 24-hour ECG test was performed on each patient utilising the pseudo-orthogonal lead configurations (X, Y, and Z).

From eligible individuals, a 24-hour ECG Holter recording was obtained and sampled with 200 Hz. In this study, patients with missing recordings of an hour or more were excluded. Thus, the dataset included only 229 patients; split into 105 HFpEF, 60 HFmEF, and 64 HFrfEF according to the ASE/EACVI guidelines. More details on the selected dataset are provided in Table 1.

Pre-processing and feature extraction

The 24-hour circadian ECG was fixed to start at 12:00 am for all the patients using a cosinor fitting analysis [32]. The SDRM-ADF filter [33] was used to filter each hour of the Holter ECG data for noise. Time- and frequency-domain features were extracted from the denoised

Table 2. Definitions of ECG features.

| ECG Feature | Definition |
|------------------------|---|
| TP (ms) | Signifying the period when the ventricles are repolarized and at rest. Variations in the TP interval may show abnormalities in repolarization, which can be linked to impaired cardiac function in HF patients. |
| QT (ms) | Reflects the total time for ventricular electrical activity, including both depolarization and repolarization. It is influenced by heart rate and can be used to assess the risk of arrhythmias. |
| ST (ms) | Reflects the period when the ventricles are electrically inactive and transitioning from depolarization to repolarization. Deviations in the ST segment can indicate myocardial ischemia or injury. |
| QRS (ms) | Represents the contraction of the ventricles, responsible for pumping blood to the systemic and pulmonary circulation. |
| Mean-Entropy | Used to analyse the complexity and unpredictability of the ECG. Higher mean entropy might indicate more irregular and complex patterns, which could be linked to various conditions or changes in the cardiac function. |
| Mean-Instant Frequency | A measure of how rapidly the frequency of a signal is changing at a particular point in time. Changes in instant frequency could relate to heart rate variability or arrhythmias. |

<https://doi.org/10.1371/journal.pone.0295653.t002>

ECGs, as explained next. The P-QRS-T components of the ECG wave have been accurately analysed using a variety of intricate approaches. A typical ECG is composed of the P wave, QRS complex, and T wave. Electric currents created by the ventricles' depolarization before to contraction, through the ventricular myocardium's depolarization extension, result in the QRS complex. The P wave, on the other hand, is created by electrical currents that the atria depolarize before contracting. The related definition and description of the ECG features are provided in Table 2.

The QT interval varies significantly according to gender, age, heart rate, and drug use. As a result, the QT interval can be calculated using Bazett's method while taking variations into account [34]. RR refers for the R-R interval, while QTc ($QTc = QT/RR$) stands for the QT interval corrected. The Slope Intersect (SI) approach was used in this study's automatic QT interval detection [35]. The Pan Tompkins ECG QRS detector was applied in Matlab to find the QRS complex [36]. Furthermore, TP, PR, and ST-T were extracted using the ECG Matlab Toolbox [37]. The power spectrogram's initial moment was used to estimate the instantaneous frequency of the signal using 255-time frames. In addition, the spectral entropy of the power spectrogram was calculated to gauge how flat and spiky a signal's spectrum is.

Machine learning and training settings

To provide an effective regression process, it is crucial to have the ability to train and estimate LVEF levels in HF patients using machine learning algorithms. In this study, four models were used, including Support Vector Machine (SVM), Extreme Gradient Boosting (XGBOOST), Gaussian Process Regression (GPR), and Decision Tree Regression (TREE), to assess the effectiveness of various machine learning methods.

1) SVM

A standard and adaptable machine learning method is often used for classification problems. However, Support Vector Regression (SVR), a variation of SVM, can be utilised for regression tasks. It manages interactions between the input characteristics and the target variable that are both linear and nonlinear. It captures complicated patterns and nonlinearity by mapping the data into a higher-dimensional space using a kernel function [38]. SVR is less sensitive to outliers compared to some other regression techniques.

Finding a "tube" or margin around the best-fitting function is the fundamental tenet of SVR; data points inside the tube are regarded as well-fitted, while those outside the tube are penalised. Using a margin around the regression line allows for some error tolerance leading to robustness to outliers.

SVM regression has only a few tuning parameters, such as the choice of kernel and the regularization parameter, making it easy to use and implement [39]. The two main hyperparameters to tune are the kernel and the regularization parameter;

1. Kernel Selection: SVRs can capture nonlinear relationships since it uses a kernel function to translate the input into a higher-dimensional space. The SVR model's performance can be considerably impacted by kernel selection. Typical kernels include linear kernel, polynomial kernel, and radial basis function (RBF).

2. Regularization Parameter (C): The regularization parameter (C) determines the trade-off between maximizing the margin and minimizing the training error. A smaller C allows a larger margin but may lead to more training errors, while a larger C leads to fewer errors but may result in a smaller margin.

In this work, the kernel types are selected to be Linear, Polynomial, and RBF, and we defined the C Values in the range of: [0.1, 1, 10,14].

2) XGBOOST

It is an ensemble learning technique that builds a powerful predictive model by combining several weak predictive models, often decision trees, commonly used to predict heart diseases [40, 41]. The XGBoost approach is a development of gradient boosting, adding decision trees to the model iteratively while each tree tries to fix the mistakes caused by the preceding one. However, XGBoost incorporates several enhancements to improve performance, speed, and generalization; to reduce overfitting and increase model generalisation, this model incorporates L1 (Lasso) and L2 (Ridge) regularisation terms. In addition, it employs a novel method to accelerate convergence while improving model accuracy by optimising the loss function during tree construction.

XGBoost provides a measure of feature importance, which helps find the most noteworthy features for making predictions. It can automatically handle missing data in the input features, making it more robust in real-world datasets. It also supports k-fold cross-validation, aiding in assessing model performance and hyperparameter tuning. Although many hyper-parameters are included in XGBoost, in this work, the hyper-parameters adopted for tuning include:

- 1) Learning Rate: The step size shrinkage used in the update to prevent overfitting. Lower values make the boosting process more conservative.

- 2) Maximum Depth: The maximum depth of a tree. Deeper trees can capture more complex patterns but may lead to overfitting.

- 2) Gamma: The minimum loss reduction required to partition a leaf node further. It acts as a regularization term, controlling the complexity of the trees and preventing overfitting.

The learning rate (η) is set to be in the range of: [0.1, 0.2], the maximum depth values are: [3, 5], and the selected gamma values are: [0, 0.01]. We use the default settings in Matlab for the rest of the other hyperparameters.

3) GPR

A potent non-parametric probabilistic regression method for modelling and forecasting continuous data is called GPR. It is based on Gaussian processes [42], which are groups of randomly distributed variables with a common distribution. In GPR, the objective is to develop a distribution over potential functions that could explain the data to represent the link between

input data and output values. In contrast to parametric regression models, which require learning specific model parameters, GPR predicts a distribution across functions consistent with the training data. A mean function and a covariance function, referred to as a kernel function, represent this distribution. The fundamental assumption in GPR is that nearby data points in the input space should have similar output values. The covariance function captures the similarity between data points and plays a crucial role in shaping the posterior distribution over functions. The choice of covariance function determines the smoothness and complexity of the resulting regression model.

GP has also been used in cardiovascular modelling [43–45] as a machine learning tool for various regression tasks. It is particularly useful when dealing with small to moderate-sized datasets, where it can capture complex patterns and provide uncertainty estimates in predictions. The hyperparameters in GPR can be broadly categorized into two types: kernel hyperparameters and regularization hyperparameters.

1) Kernel Function: The choice of the kernel function, also known as the covariance function or radial basis function, determines how the covariance between data points is computed. Common kernels include the radial basis function (RBF or squared exponential), Matérn, polynomial, and more.

2) Regularization Hyperparameters (Alpha): In GPR implementations, regularization hyperparameters may be available to prevent overfitting and improve model generalization. These hyperparameters control the trade-off between fitting the training data and simplifying the model.

The kernel functions in this work are set to Rational Quadratic and Matern52. In addition, the alpha range is selected to be: [0.01, 0.1, 1].

4) TREE

Decision Tree Regression is a supervised machine-learning technique used for regression tasks. Decision tree regression predicts continuous numerical values unlike decision trees for classification, which output discrete class labels. The core idea behind decision tree regression is to partition the feature space into regions and assign a constant value to each region, which predicts any input falling into that region.

The algorithm begins by partitioning the input data into subsets based on the values of the input features. It selects the feature and the corresponding threshold that best splits the data, aiming to minimize the variance (or other suitable loss function) of the target values within each partition. The selection is based on the chosen feature and threshold at each internal node of the tree (the decision node) [46]. The algorithm determines if the feature value of the input data point exceeds or falls short of the threshold. Depending on the result, the algorithm moves to the left or right child node.

Each subset is subjected to data partitioning and decision-making, resulting in a hierarchical structure of decision and leaf nodes. The target variable is given a prediction value when the algorithm reaches a leaf node (terminal node). The mean or median of the goal values for that location serves as this prediction value. Based on the feature values of the input, the algorithm moves through the tree from the root node to a particular leaf node to forecast a new input sample. The model's output for that input is the predicted value at the leaf node.

The interpretability of Decision Tree Regression is its main benefit because the output tree can be inspected and understood immediately. It is robust against outliers and can handle numerical and categorical data, making it widely used in heart disease modelling and prediction [47]. However, obtaining the best performance and reliable generalisation requires rigorous hyperparameter altering and countermeasures against overfitting. In the context of tuning

decision tree hyperparameters, several key parameters are considered, focusing on controlling the tree's complexity and potential for overfitting. These hyperparameters include but are not limited to maximum depth, minimum samples split, minimum samples leaf, and maximum leaf nodes. During the hyperparameter tuning process, the Min Leaf Sizes are set to: [2, 4, 6, 8, 10], and the Surrogate Options are chosen from the set: [off, on]

The best SVR hyperparameter combination was $C = 14$ and a linear kernel type. A high C value indicates that the SVR model has a low tolerance for errors in the training data. In other words, the model aims to fit the training data as closely as possible, even if it means allowing more margin violations. Using a linear kernel, the SVR model attempts to draw a hyperplane in the feature space that best fits the data points while maximizing the margin between the positive and negative support vectors. This simplicity and interpretability of the linear model make it advantageous in scenarios where the relationship between the input features and the target variable can be adequately approximated by a linear function [48]. For XGBoost, the best performance was achieved with a Learning Rate of 0.1, a Maximum Depth of 3, and a Gamma value of 0.01. A moderate Learning Rate, a shallow Maximum Depth, and a conservative Gamma value indicate a well-balanced XGBoost model. It leverages a cautious approach in building trees to avoid overfitting while still allowing the model to capture relevant data patterns in the data effectively [49]. This hyperparameter combination will likely produce a robust and accurate XGBoost model, providing reliable estimates of LVEF levels in heart failure patients across distinct categories (HFpEF, HFmEF, and HFrEF).

In the case of GPR, the best performance was observed using the kernel function of rational quadratic and an alpha of 0.1. The rational quadratic kernel combines the characteristics of both the squared exponential kernel and the Matérn kernel. Matérn kernel is a versatile covariance function that can adapt to various levels of smoothness in data and provides a balance between the squared exponential kernel and the Gaussian kernel. The rational quadratic kernel is more flexible than the squared exponential kernel, allowing it to capture various patterns in the data, including short-range and long-range correlations. An alpha value of 0.1 indicates that the model is giving more importance to the squared exponential term, making it more sensitive to large-scale variations [42]. This choice could be appropriate if the underlying data has smooth patterns and long-range dependencies. Finally, for the Decision Tree Regression model, the best hyperparameter combination comprised a Min Leaf Size of 4 and 'off' for the Surrogate Option. Choosing a minimum leaf size of 4 strikes a balance between capturing some of the complexity in the data while preventing excessive overfitting [50]. This choice suggests that the data might have sufficient complexity to warrant more splits but still benefits from regularization to control overfitting. Our results indicate that each regression model exhibited exceptional performance when tuned with its best hyperparameter combination.

Training and testing configuration

This study introduces a nested loops framework (Fig 1), which synergistically combines leave-one-out cross-validation (LOOCV) as the outer loop and 5-fold cross-validation as the inner loop. This carefully designed approach is advocated for its ability to offer a robust and comprehensive methodology for model evaluation and hyperparameter tuning. In the outer loop, each data point is iteratively withheld as a validation set while the model is trained on the remaining data instances. This process is repeated for all data points, allowing comprehensive evaluation across various training and validation scenarios to obtain reliable performance estimates. Within the outer loop, the inner loop performs 5-fold cross-validation, partitioning the data into five subsets. The model is trained on four subsets while the remaining one is validated. This cycle is repeated five times, robustly assessing the model's performance. The

average performance metrics from these iterations are then utilized to evaluate the model's effectiveness more comprehensively.

To optimize the model's hyperparameters efficiently, the nested loop approach incorporates grid search technology. Grid search systematically explores a predefined hyperparameter grid, covering various combinations of hyperparameter values. At each iteration of the nested loops, the model's performance is evaluated based on different hyperparameter configurations. This allows for an exhaustive search for the optimal set of hyperparameters that maximizes the model's accuracy and generalization capabilities.

Using nested loops with LOOCV, 5-fold cross-validation, and grid search technique provides a rigorous and reliable framework for regression tasks in machine learning. It enables meticulous hyperparameter tuning, ensures robust model evaluation, and ultimately leads to the selection of the best-performing regression model with enhanced predictive capabilities of the LVEF estimation and increased applicability to heart failure diagnosis.

The evaluation of the regression model's performance using key metrics, including the average Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and correlation analysis, were computed. Furthermore, the estimation process was visualized through Bland-Altman plots and correlation plots [48, 51], comprehensively analysing the model's predictive capabilities.

Furthermore, the feature importance metric quantifies the impact of each feature on reducing the impurity (e.g., RMSE) in the model. The more a feature contributes to reducing the impurity during tree construction, the more important it is considered. This importance is typically computed as the total reduction in impurity achieved by each feature over all nodes in the tree. In nested cross-validation (LOOCV and 5-fold cross-validation), the feature importance is calculated iteratively over multiple folds or model training and evaluation iterations. This ensures robustness in estimating feature importance and helps mitigate overfitting issues.

Results

The single GeForce NVIDIA MX350 (2 GB) memory, Intel Core i5- 1135G7 (11th Gen) used to train and validate the models hourly. The training and testing phases took 10–20 minutes per model due to following the nested loop scheme. After tuning each model, the optimal hyperparameter combinations for each model are shown in Table 3.

All models were trained using the extracted ECG features (TP, QTc, PR, ST-T, QRS, Mean-Entropy, and Mean-Infrequency), Table 4 summarizes the values of these features for the three groups (HFpEF, HFmEF, and HFREF) with statistical comparisons. Table 5 shows the average RMSE values for the estimated LVEF levels, which vary over 24 hours and across different regression models with the best hyperparameters combination. In this table, each column corresponds to a specific time throughout the 24-hour cycle, and each row stands for a different

Table 3. Regression models and hyperparameters tuning.

| Regression Model | Hyperparameters | Best Combination |
|------------------|---|---|
| SVM | C Values: [0.1, 1, 10,14] Kernel Types: [linear, polynomial, rbf]. | C Value = 14 Kernel Type = linear |
| XGBOOST | Learning Rate: [0.1, 0.2] Maximum Depth: [3, 5] Gamma: [0, 0.01] | Learning Rate = 0.1 Maximum Depth = 3 Gamma = 0.01 |
| GPR | Kernel Functions: [rational quadratic, matern52] Alpha Range: [0.01, 0.1, 1] | Kernel Function = rational quadratic Alpha Range = 0.1 |
| TREE | Min Leaf Sizes: [2, 4, 6, 8, 10] Surrogate Options:[off, on] | Min Leaf Size = 4 Surrogate Option = off |

<https://doi.org/10.1371/journal.pone.0295653.t003>

Table 4. ECG features for the three groups (HFpEF, HFmEF, and HFrEF).

| ECG Feature (\pm SD) | HFpEF | HFmEF | HFrEF | p-Value |
|-------------------------|------------------|------------------|------------------|---------|
| TP-Interval (ms) | 198 \pm 23.5 | 200 \pm 31.3 | 276 \pm 35.2 | 0.0014 |
| QTc-Interval (ms) | 451 \pm 54.1 | 456 \pm 44.2 | 501 \pm 33.7 | <0.0001 |
| ST-T Interval (ms) | 302 \pm 53.6 | 251 \pm 46.7 | 248 \pm 38.9 | 0.0007 |
| QRS-Interval (ms) | 96 \pm 18.3 | 102 \pm 31.1 | 117 \pm 28.5 | <0.0001 |
| Mean-Entropy | 0.73 \pm 0.022 | 0.83 \pm 0.042 | 0.76 \pm 0.036 | 0.0015 |
| Mean-InsFreq | 38.4 \pm 8.7 | 36.1 \pm 10.5 | 36.3 \pm 11.2 | 0.0031 |

<https://doi.org/10.1371/journal.pone.0295653.t004>

regression model with the best hyperparameters combination. The RMSE values indicate the average magnitude of the errors between the estimated LVEF levels and the actual values for each period and model. Lower RMSE values show better model performance, indicating smaller errors and more accurate predictions. The table provides an overview of how the performance of the different regression models varies across the 24 hours. It allows for comparing different models and their suitability for estimating LVEF levels at contrasting times of the day.

The lowest error (RMSE = 3.4) occurred between 10 pm to 11 pm using the Tree function, where the minimum leaf size is 4, and the surrogate way is set to be off. While the GPR function led to an RMSE value of 3.8 at this time interval, with rational quadratic kernel function and $\alpha = 0.1$. The XGBOOST best performance occurs in the morning interval (8–9 am) with an RMSE value of 4.1. The early morning time (12–1 am) results in the lowest error using the SVM model with a linear function and C value of 14. In addition, Fig 2 depicts the RMSE distribution across the 24 hours using the four regression models. In the latter, the time interval of the lowest possible RMSE values is highlighted in red, where the Tree regression model outperforms the other models in the three-time intervals.

Furthermore, Fig 3 shows the original and estimated LVEF values for all patients at 00:00–01:00 h, 08:00–09:00 h, and 22:00–23:00 h with the lowest RMSE values. The dashed lines represent the ranges of the heart failure patients with reduced, midrange, and preserved LVEF according to ASE/EACVI guidelines. To elaborate more on the prediction models, Figs 4 and 5 show the correlation and Bland-Altman plots, respectively. The highest correlation coefficient occurred at 10–11 pm using the Tree model with 0.9227, while the correlation at 10–11 pm and 8–9 am was 0.9139 and 0.8920 using the GPR and XGBOOST, respectively. And the least correlation was from midnight to 1 am, using the SVM model. Furthermore, the mean bias was overestimated at 10–11 pm and 8–9 am while underestimated between midnight and 1 am.

Moreover, in this study, we have evaluated the feature importance using the best regression model. The results are presented in Fig 6, highlighting the relative significance of each feature in contributing to the model's predictive performance. The feature importance values have been calculated based on their influence on the target variable, providing valuable insights into the key determinants driving the LVEF estimation. This analysis allows us to identify the most influential features, enabling a deeper understanding of the underlying factors influencing the outcome of the regression model. The comprehensive evaluation of feature importance aids in interpreting the model's behaviour and helps informed decision-making for our research application.

Furthermore, analyzing RMSE values through the 24 hours allows us to assess how sensitive the models are to circadian changes at different times of the day. If the RMSE values exhibit significant fluctuations or patterns, it suggests that the model is sensitive to circadian changes. As shown in Fig 7, the p-values in the correlation matrix indicate the statistical significance of

Table 5. Average RMSE values through 24-Hours, using regression models with best hyperparameters.

| Regression Model | RMSE | | | | | | | | | | | | | | | | | | | | | | | |
|------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 00:01 | 01:02 | 02:03 | 03:04 | 04:05 | 05:06 | 06:07 | 07:08 | 08:09 | 09:10 | 10:11 | 11:12 | 12:13 | 13:14 | 14:15 | 15:16 | 16:17 | 17:18 | 18:19 | 19:20 | 20:21 | 21:22 | 22:23 | 23:24 |
| SVM | 4.2 | 4.5 | 5.2 | 4.8 | 4.6 | 5.8 | 4.9 | 4.7 | 4.8 | 5.0 | 4.6 | 5.3 | 4.7 | 5.0 | 4.3 | 4.7 | 4.4 | 4.9 | 5.1 | 4.4 | 4.7 | 4.7 | 4.5 | 5.2 |
| XGBOOST | 4.4 | 4.3 | 5.0 | 4.4 | 4.5 | 5.6 | 4.9 | 4.6 | 4.1 | 4.9 | 4.5 | 5.2 | 4.7 | 5.0 | 4.3 | 4.8 | 4.3 | 4.9 | 5.0 | 4.2 | 4.2 | 4.3 | 4.2 | 5.0 |
| GPR | 3.9 | 4.6 | 4.4 | 5.1 | 4.1 | 4.7 | 5.4 | 5.1 | 4.2 | 5.0 | 4.9 | 4.5 | 5.1 | 4.8 | 5.3 | 4.3 | 4.9 | 4.4 | 4.7 | 4.9 | 4.3 | 4.3 | 3.8 | 4.5 |
| Tree | 3.6 | 4.4 | 4.3 | 5.0 | 4.2 | 4.5 | 5.5 | 4.9 | 4.0 | 4.7 | 4.9 | 4.9 | 5.5 | 4.8 | 5.1 | 5.7 | 4.8 | 4.3 | 4.9 | 4.9 | 4.2 | 4.2 | 3.4 | 4.3 |

Yellow shading stands for the lowest RMSE value

<https://doi.org/10.1371/journal.pone.0295653.t005>

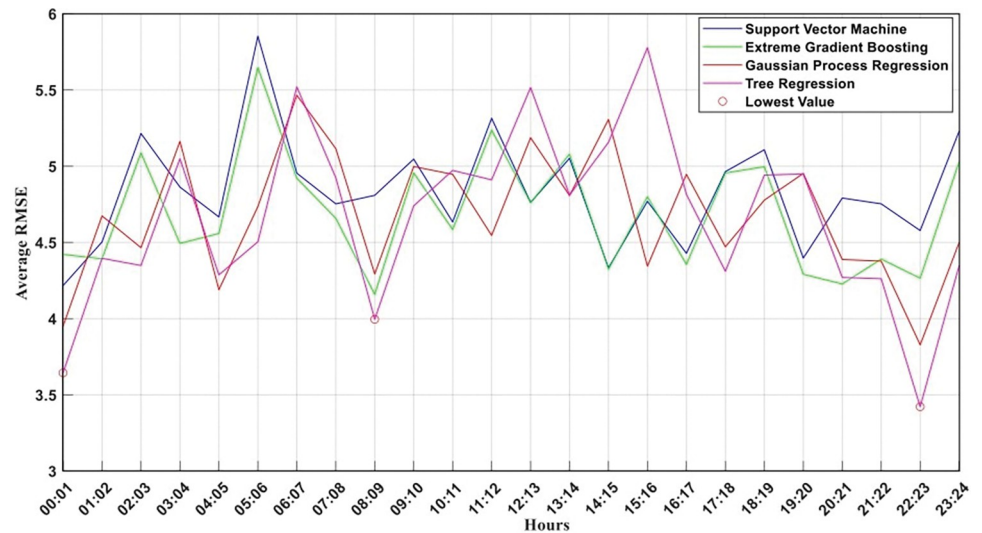


Fig 2. RMSE per hour using the best hyperparameters combination for the four regression models. Red circles show the hours of occurrence of the lowest RMSE values. Where QTc and QRS are found to be the most important features for evaluating the LVEF levels.

<https://doi.org/10.1371/journal.pone.0295653.g002>

the observed correlations using the best regression model. A low p-value suggests that the correlations are unlikely to be due to random chance, reinforcing the validity of the relationships. Whereas, the model is sensitive to circadian changes when moving from morning to afternoon hours (e.g., at 1–12 h, 2–13 h, 7–12 h and 8–14 h). Also, this analysis demonstrates the model's ability to capture circadian changes at the afternoon intervals (e.g., 13–17 h, 13–18 h, and 15–18 h), and at the evening intervals (e.g., 18–23 h, and 22–24 h).

Discussion

This study provides valuable insights into the importance of utilizing ECG features to estimate LVEF levels in heart failure patients categorized into HFpEF, HFmEF, and HFrEF groups. Through a comprehensive investigation, the study identifies the significance of specific ECG features in accurately predicting LVEF levels throughout a 24-hour study.

By employing regression analysis, the study focuses on minimizing the RMSE to achieve the best possible predictive performance. By identifying the optimal hyperparameters and model configurations, the study successfully achieves high accuracy in each of the four regression schemes at various hours in the heart's circadian rhythm. This remarkable performance demonstrates the potential of ECG features integrated into machine learning algorithms as a valuable assistive screening tool for heart failure diagnosis and treatment. These ECG parameters, such as QRS duration, ST segment changes, and T-wave abnormalities, closely mirror the temporal aspects of ventricular systole and diastole. The QTc interval, representative of ventricular repolarization, intricately mirrors the duration required for myocardial cells to recover their resting state following depolarization. As LVEF signifies the efficacy of ventricular contraction, its prediction is inherently linked to the temporal aspects of cardiac repolarization reflected in the QTc interval. Similarly, the QRS duration reflects the propagation of electrical impulses during ventricular depolarization, intricately tied to mechanical events. Changes in QRS duration can be indicative of altered ventricular conduction, which in turn influences the synchronization of contraction, affecting LVEF. Thus, by harnessing the electrophysiological nuances encoded in ECG features, the models capitalize on the profound connection between

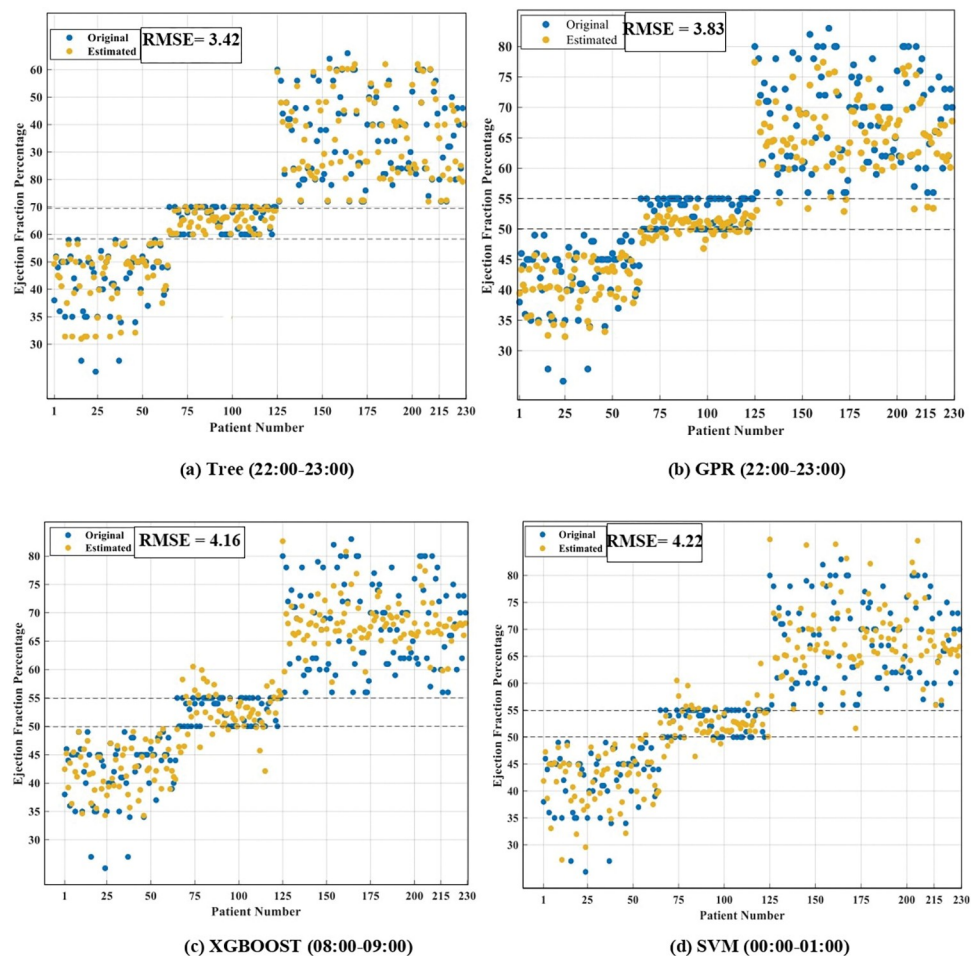


Fig 3. Original and estimated LVEF values per patient for hours (22:00–23:00, 08:00–09:00, 00:00–01:00) using the Tree, GPR, XGBOOST and SVM, respectively.

<https://doi.org/10.1371/journal.pone.0295653.g003>

electrical signalling and mechanical performance. This synergy offers a more comprehensive understanding of cardiac function, rendering the models adept at predicting LVEF with heightened accuracy, drawing from the intricate dialogue between electrophysiological dynamics and mechanical behaviours.

Compared to previous work (Table 6), SVR models were employed to predict LVEF using heart rate variability (HRV) data derived from 24-hour ECG recordings of 92 patients participating in the IDEAL study [52]. The patients were categorized based on their LVEF as preserved, mid-range, or reduced. The SVR models exhibited varying performance throughout the 24 hours. The most accurate estimation, with an RMSE of 10.4, was achieved during the time interval between 6 pm and 7 pm. The polynomial kernel function was employed for this interval, and the model incorporated 18 of the 25 available features. In another study [53], three different models, namely SVM with RBF kernel, Generalized Linear Model (GLM), and CNN (Convolutional Neural Network), were employed to estimate the LVEF in heart failure patients belonging to the three distinct categories. The study used clinical profiles from 303 patients with CAD acquired from American and Greek patient databases. The goal was to compare the performance of the three models in predicting LVEF in these patients. After rigorous evaluation, it was observed that the CNN model surpassed the other approaches with an RMSE of 4.13 and a correlation coefficient of 0.85.

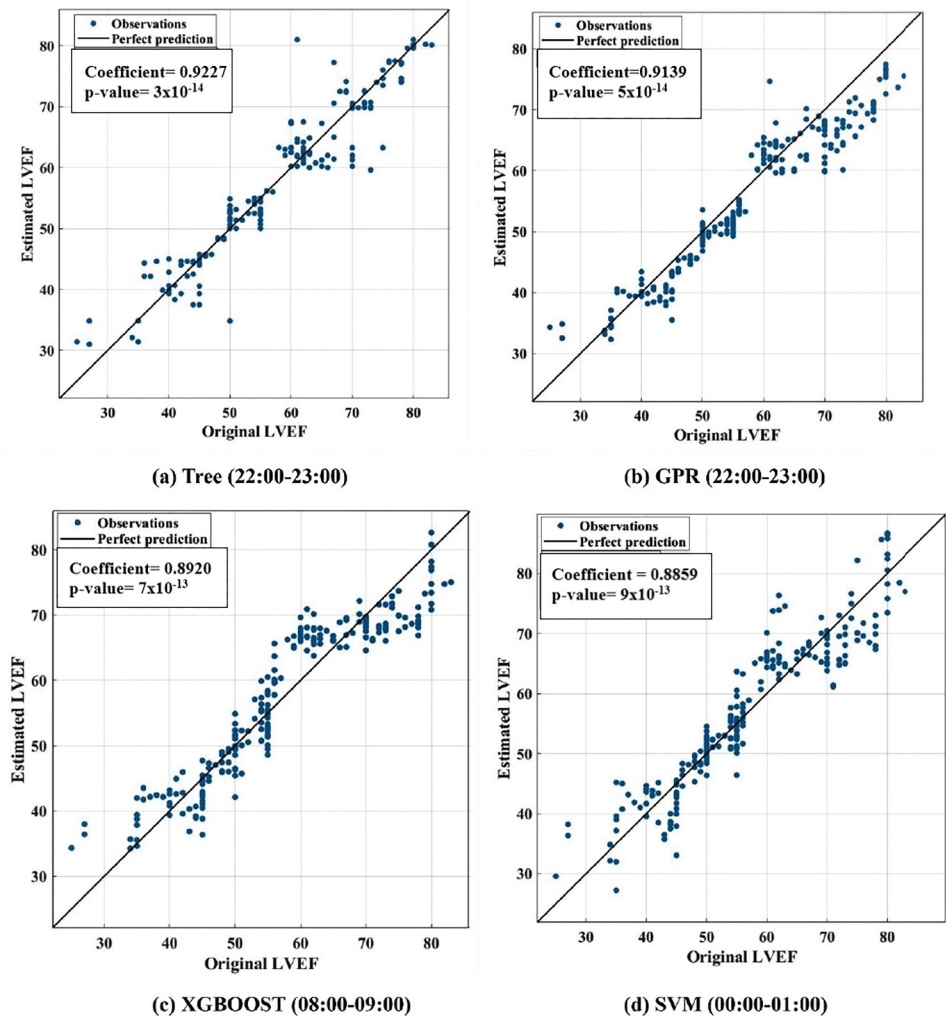


Fig 4. Correlation plots between the original and estimated LVEF values for hours (22:00–23:00, 08:00–09:00, 00:00–01:00) using the Tree, GPR, XGBOOST and SVM, respectively.

<https://doi.org/10.1371/journal.pone.0295653.g004>

The present study surpasses the performance of previous works [52–54] in the literature for estimating LVEF levels for heart failure patients across the three distinct categories. The results of this investigation reveal that the optimal performance is achieved during the time interval between 10 pm and 11 pm, using a tree-based model. This model exhibits an impressive RMSE of 3.42, a high correlation coefficient of 0.92, and a Bland-Altman analysis demonstrating a mean difference of 0.16 with a narrow limit of agreement of $\pm 6.85\%$. Furthermore, the MAE is remarkably low at 2.28%, meaning that the model's predictions are closer to the actual values.

Clinical relevance and ECG features importance

ECG features obtained through ambulatory Holter monitoring have been established as valuable predictors of total mortality and the progression of heart failure. This continuous monitoring technique offers unique insights into a patient's cardiac health, allowing for a comprehensive evaluation of heart rhythm and electrical activity over an extended period.

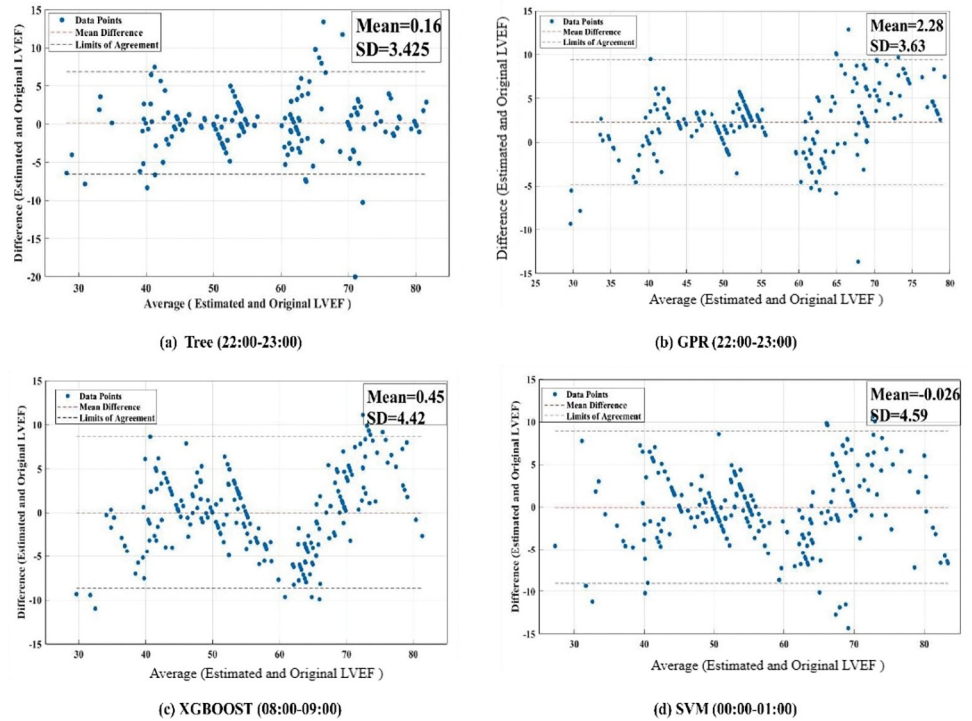


Fig 5. Bland-Altman plots between the average value of (Original and estimated LVEF) and their corresponding difference for hours (22:00–23:00, 08:00–09:00, 00:00–01:00) using the Tree, GPR, XGBOOST and SVM, respectively.

<https://doi.org/10.1371/journal.pone.0295653.g005>

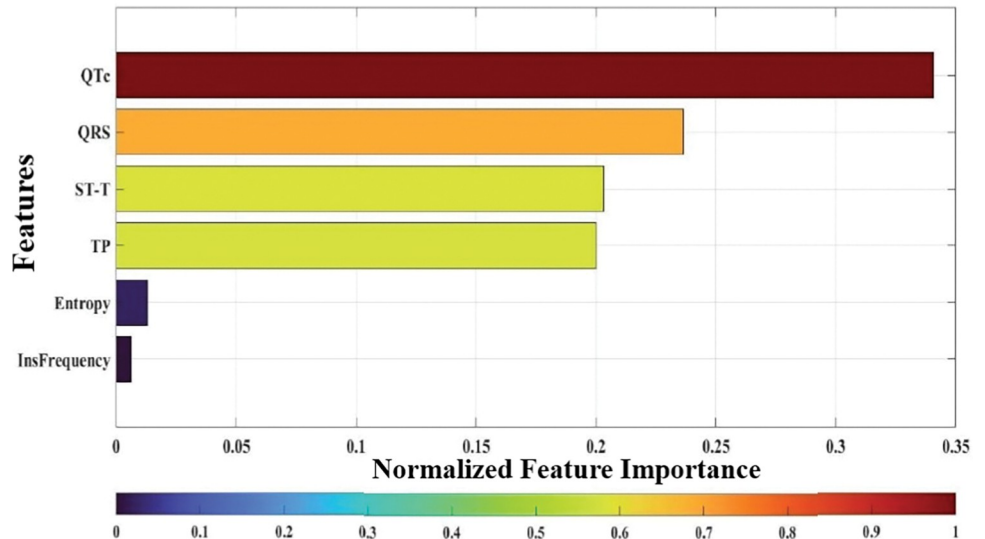


Fig 6. Feature importance over the Tree regression model (10:00–11:00pm). This study used normalized importance scores for the ECG features (QTc, QRS, ST-T, TP, Entropy, and Instant Frequency) to estimate the LVEF levels in the three HF categories. Indicating that QTc is the most important feature for evaluating the LVEF levels.

<https://doi.org/10.1371/journal.pone.0295653.g006>

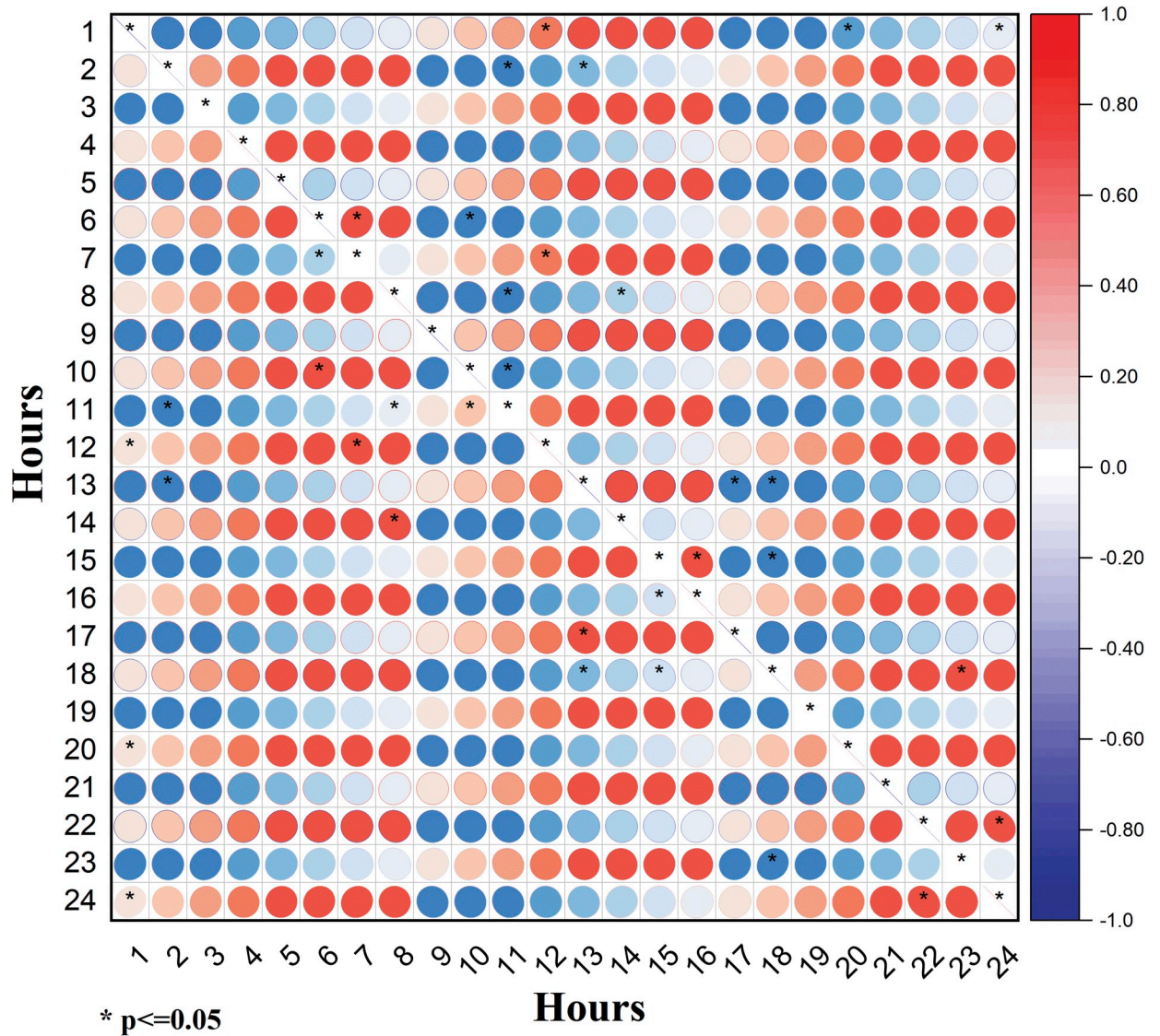


Fig 7. Correlation matrix of the RMSE values through 24-Hours, using the best regression model. Black asterisks show the hours of statistical significance to circadian changes.

<https://doi.org/10.1371/journal.pone.0295653.g007>

Table 6. Summary of the studies for LVEF estimation in preserved, midrange, and reduced HF patients. Best performance highlighted in yellow.

| Author (Year) | ECG Signal Duration | Data | Method/Features | Regression | RMSE | Correlation Coefficient | Bland-Altman Mean \pm 2 Std | MAE |
|----------------------|--------------------------|--------------------------------------|---------------------|------------|------|-------------------------|-------------------------------|------|
| 54 | 1hour (24h study) | 92 (American Cohort) | HRV features | SVR | 10.4 | 0.61 | 0.53 \pm 20.44 | NA |
| 55 | NA | 303 (American, Greek Cohorts) | Clinical profiles | CNN | 4.13 | 0.85 | 0.39 \pm 11.61 | NA |
| Current study | 1hour (24h study) | 229 (American, Greek Cohorts) | ECG features | SVR | 4.22 | 0.89 | -0.03 \pm 9.18 | 3.13 |
| | | | | XGBOOST | 4.16 | 0.89 | 0.45 \pm 8.84 | 3.24 |
| | | | | GPR | 3.83 | 0.91 | 2.28 \pm 7.26 | 2.94 |
| | | | | Tree | 3.42 | 0.92 | 0.16 \pm 6.85 | 2.28 |

<https://doi.org/10.1371/journal.pone.0295653.t006>

Even when accounting for other known risk factors, such as age, gender, and comorbidities, the information derived from Holter monitoring provides additional and valuable prognostic information. However, estimating LVEF in heart failure patients hourly using ECG features can offer valuable insights into the temporal patterns and fluctuations of cardiac function. This hourly analysis can enhance our understanding of how LVEF levels vary throughout the day, potentially identifying when LVEF estimation is more accurate and best correlated with clinical variations or when heart function is at its optimal state. This understanding of temporal variation may be crucial for adjusting treatment strategies or interventions based on the patient's circadian rhythm and physiological fluctuations. This work demonstrates superior performance during the evening (22–23) and early morning hours (12 am–1 and 8–9) for estimating LVEF levels. Additionally, these hourly intervals align with previously reported high-risk periods for increased mortality and myocardial infarcts observed in the early morning and evening hours [55–57].

Furthermore, using the most effective regression model, we evaluated feature importance during this study. Our findings revealed that the most crucial features contributing to LVEF prediction are QTc, QRS, ST, and TP in descending order of importance. These specific ECG parameters significantly influenced the target variable, supplying valuable insights in characterizing cardiac function and highlighting their potential as important clinical indicators in assessing heart failure patients.

The prominence of QTc as the most important feature can be attributed to several key factors [58–61], QTc interval stands for the duration of ventricular depolarization and repolarization, encompassing the electrical conduction time within the heart. In HF patients, electrical conduction abnormalities are common due to ventricular remodelling and changes in ion channel kinetics. Prolonged QTc intervals may show delayed repolarization, which can be associated with impaired ventricle (which involves alterations in ventricular size, shape, and function) and reduced LVEF. HF patients are at increased risk of arrhythmias, and QTc prolongation is a known risk factor for life-threatening ventricular arrhythmias, abnormal QTc intervals are indicative of arrhythmogenic substrate and can influence LVEF by potentially triggering life-threatening arrhythmias that may affect cardiac function. In addition, QTc interval is influenced by autonomic nervous system activity, specifically parasympathetic and sympathetic inputs to the heart. HF patients often have altered autonomic regulation, and QTc prolongation may be linked to autonomic dysfunction, which can affect cardiac function and LVEF. Some medications commonly used in HF management can affect the QTc interval. The QTc-prolonging effects of certain drugs may contribute to the importance of this feature in the model, as medication use is an essential consideration in LVEF estimation.

Several studies have shown that QRS duration is a prognostic marker in HF patients [62–64]. The QRS complex is the ventricles' electrical activation and later mechanical contraction. In HF patients, mechanical desynchrony is a common phenomenon, where different regions of the ventricles contract uncoordinated. QRS prolongation can indicate bundle branch blocks (BBBs). These conduction abnormalities can affect the ventricular function and contribute to reduced LVEF [65]. In addition, for patients undergoing cardiac resynchronization therapy (CRT), QRS duration is a critical parameter for determining eligibility for CRT implantation. CRT effectively improves mechanical synchrony in HF patients with electrical desynchrony and prolonged QRS duration. By considering QRS duration in LVEF prediction, clinicians can gain insights into potential CRT candidacy and tailor treatment options for better patient outcomes.

Moreover, the STT interval is closely related to T-wave morphology. T-wave abnormalities, such as inversion or flattening, are commonly seen in heart failure patients [66–68]. These T-wave changes can indicate valuable information about myocardial ischemia or injury and can

contribute to LVEF prediction accuracy. Variations in the TP interval may show abnormalities in repolarization, which can be linked to impaired cardiac function in HF patients. Those alterations have been associated with an increased risk of arrhythmias, such as ventricular tachycardia and ventricular fibrillation. Several studies have reported a correlation between abnormal TP intervals and reduced LVEF in HF patients. Prolonged TP intervals have been associated with impaired ventricular function, while shortened TP intervals may show myocardial remodelling and potential deterioration of cardiac performance.

The prioritization of these specific features based on their impact on the LVEF estimation reinforces the model's ability to discern critical patterns and relationships among various ECG parameters. This knowledge is instrumental in enhancing our understanding of the underlying physiological mechanisms that govern cardiac health and dysfunction.

The proposed technology, which leverages regression models based on circadian ECG features, has the potential to serve as a valuable tool for early HF diagnosis and risk assessment. A more detailed explanation of its practical applications and requirements for implementation:

1. Data collection and Integration:

- Data sources: Access to ECG data is a fundamental requirement. This data can be collected from various sources, including hospitals, clinics, and ambulatory ECG monitors.
- EHR Integration: Integration with electronic health records (EHR) systems is essential for streamlined data access and patient history analysis. This enables healthcare providers to have a comprehensive view of the patient's health status.

2. Machine learning infrastructure:

- Training and validation: Rigorous model training and validation, using large and diverse datasets, are critical to ensure the reliability and accuracy of the technology.

3. Clinical validation:

- Collaboration with healthcare professionals: Close collaboration with cardiologists and other healthcare professionals is crucial for the clinical validation of the technology. Their expertise ensures that the technology aligns with established medical practices and standards.
- Clinical trials: Conducting clinical trials to assess the technology's performance in real-world healthcare settings is a key component. This helps establish its clinical efficacy and safety.

4. Regulatory compliance:

- Quality assurance: Implementing quality assurance protocols and adhering to Good Clinical Practice (GCP) and Good Manufacturing Practice (GMP) standards are important for maintaining high-quality healthcare technology.

5. Privacy and security:

- Data protection: Safeguarding patient data is paramount. Adherence to data protection regulations, such as Health Insurance Portability and Accountability Act (HIPAA) in the United States or General Data Protection Regulation (GDPR) in the European Union, is mandatory.
- Security measures: Implementing robust security measures to protect patient data from breaches and unauthorized access is essential.

6. Clinical workflow:

- Integration into ECG testing: Our technology can be seamlessly integrated into the clinical workflow as a supplementary tool during routine ECG testing.
- Automated LVEF estimation: When an ECG is performed, the technology can automatically generate an LVEF estimation, which is then made available to the attending physician.
- Physician assessment: Physicians can use this information to identify patients at risk for HF and initiate further diagnostic tests or interventions as necessary.

In conclusion, implementation in medical practice involves adherence to specific protocols related to data collection, machine learning infrastructure, clinical validation, regulatory compliance, data privacy, and a secure clinical workflow. These protocols are vital to ensure the accurate and ethical application of this technology for the benefit of patients and healthcare providers.

Limitations and future work

Despite the effectiveness of machine learning-based models in predicting LVEF, our study has identified several limitations associated with these models. First, we focused on using six features extracted from the circadian ECG for LVEF prediction. Nevertheless, there is a need for further exploration and investigation of more features, such as QRS area, the amplitude of peak changes, and the association with EDR (ECG-derived respiration), to gain better insights into their impact on LVEF predictions. Furthermore, while the current study used a dataset including patients from both American and Greek populations, it is essential to subject the trained models to further testing on more diverse patient groups to ensure broader applicability and generalization of their performance. In this study, the patient cohort consists of a significantly higher proportion of male participants when compared to females; future studies may apply additional investigations of LVEF prediction on a more evenly distributed dataset between male and female subjects. In future research, the regression models developed in this study may have the potential to be adapted or extended to predict various parameters related to heart health (e.g., myocardial ischemia), their applicability to underlying diseases that can lead to heart failure would require additional research and validation.

Conclusion

This study highlights the potential of ECG features derived from 24-hour recordings as a promising alternative to the current echocardiography gold standard for finding LVEF levels in CAD patients. From a machine learning perspective, the developed approach offers valuable insights into finding ECG patterns suitable for accurately estimating LVEF in different heart failure categories (preserved, midrange, and reduced).

Furthermore, the proposed study aims to extend the application of LVEF assessment to communities where access to the required instruments is limited due to economic challenges or lack of clinical expertise. The study's outcomes hold the potential to aid the development of a model to predict the HF phenotype or track its changes during therapy, providing a versatile tool for exploring disease pathophysiology and objectively assessing therapeutic approaches in future HF patients.

Notably, the proposed machine learning model offers simplicity, efficiency, and cost-effectiveness, enabling continuous cardiac analysis and representing a viable alternative to more elaborated gold standard techniques. As research advances, using more extensive datasets and leveraging deep learning techniques will help compare the proposed model's performance to other methods. Overall, this research contributes to advancing the field of ECG-based LVEF estimation and its potential applicability, particularly in resource-constrained settings.

Author Contributions

Conceptualization: Sona M. Al Younis, Leontios J. Hadjileontiadis, Mohanad Alkhodari, Ahsan H. Khandoker.

Data curation: Mohanad Alkhodari.

Formal analysis: Sona M. Al Younis.

Investigation: Sona M. Al Younis.

Methodology: Sona M. Al Younis, Leontios J. Hadjileontiadis, Aamna M. Al Shehhi, Mohanad Alkhodari, Ahsan H. Khandoker.

Project administration: Leontios J. Hadjileontiadis, Aamna M. Al Shehhi, Cesare Stefanini, Ahsan H. Khandoker.

Resources: Mohanad Alkhodari, Stergios Soulaïdopoulos, Petros Arsenos, Ioannis Doundoulakis, Konstantinos A. Gatzoulis, Konstantinos Tsioufis.

Software: Sona M. Al Younis.

Supervision: Leontios J. Hadjileontiadis, Aamna M. Al Shehhi, Cesare Stefanini, Ahsan H. Khandoker.

Validation: Sona M. Al Younis.

Visualization: Sona M. Al Younis.

Writing – original draft: Sona M. Al Younis.

Writing – review & editing: Sona M. Al Younis, Leontios J. Hadjileontiadis, Aamna M. Al Shehhi, Cesare Stefanini, Mohanad Alkhodari, Ahsan H. Khandoker.

References

1. Curtis JP, Sokol SI, Wang Y, Rathore SS, Ko D, Jadbabaie F, et al. Heart failure: preventing disease and death worldwide. *ESC Heart Failure* vol. 1 Preprint at <https://doi.org/10.1002/ehf2.12005> (2014).
2. Ziaeian B, Fonarow GC. Epidemiology and aetiology of heart failure. *Nature Reviews Cardiology* vol. 13 Preprint at <https://doi.org/10.1038/nrcardio.2016.25> (2016).
3. Bleumink GS, Knetsch AM, Sturkenboom MCJM, Straus SMJM, Hofman A, Deckers, et al. Quantifying the heart failure epidemic: Prevalence, incidence rate, lifetime risk and prognosis of heart failure—The Rotterdam Study. *Eur Heart J* 25, (2004). <https://doi.org/10.1016/j.ehj.2004.06.038> PMID: 15351160
4. Chan YK, Tuttle C, Ball J, Teng THK, Ahamed Y, Carrington MJ, et al. Current and projected burden of heart failure in the Australian adult population: A substantive but still ill-defined major health issue. *BMC Health Serv Res* 16, (2016). <https://doi.org/10.1186/s12913-016-1748-0> PMID: 27654659
5. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Colvin MM, et al. 2017 ACC/AHA/HFSA Focused Update of the 2013 ACCF/AHA Guideline for the Management of Heart Failure: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. *Circulation* 136, e137–e161 (2017). <https://doi.org/10.1161/CIR.0000000000000509> PMID: 28455343
6. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JGF, Coats AJS, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J* 37, 2129–2200 (2016). <https://doi.org/10.1093/eurheartj/ehw128> PMID: 27206819
7. Tsutsui H, Isobe M, Ito H, Okumura K, Ono M, Kitakaze M, et al. JCS 2017/JHFS 2017 Guideline on Diagnosis and Treatment of Acute and Chronic Heart Failure — Digest Version —. *Circulation Journal* 83, 2084–2184 (2019).
8. Ramani G. V., Uber P. A. & Mehra M. R. Chronic Heart Failure: Contemporary Diagnosis and Management. *Mayo Clin Proc* 85, 180–195 (2010). <https://doi.org/10.4065/mcp.2009.0494> PMID: 20118395
9. Curtis JP, Sokol SI, Wang Y, Rathore SS, Ko DT, Jadbabaie F, et al. The association of left ventricular ejection fraction, mortality, and cause of death in stable outpatients with heart failure. *J Am Coll Cardiol* 42, 736–742 (2003). [https://doi.org/10.1016/s0735-1097\(03\)00789-7](https://doi.org/10.1016/s0735-1097(03)00789-7) PMID: 12932612
10. Rodriguez J, Voss A, Caminal P, Bayés-Genis, Giraldo B. Characterization and classification of patients with different levels of cardiac death risk by using Poincaré plot analysis.
11. Hajouli S, Ludhwani D. Heart Failure and Ejection Fraction. StatPearls. 2022 Dec 23, Available from: <https://www.ncbi.nlm.nih.gov/books/NBK553115/>.

12. Pfeffer MA, Shah AM, Borlaug BA. Heart Failure with Preserved Ejection Fraction in Perspective. *Circ Res* 124, 1598–1617 (2019). <https://doi.org/10.1161/CIRCRESAHA.119.313572> PMID: 31120821
13. Fonarow G. C. & Hsu J. J. Left Ventricular Ejection Fraction. *JACC Heart Fail* 4, 511–513 (2016).
14. Tsao CW, Lyass A, Larson MG, Cheng S, Lam CSP, Aragam JR, et al. Prognosis of Adults With Borderline Left Ventricular Ejection Fraction. *JACC Heart Fail* 4, 502–510 (2016). <https://doi.org/10.1016/j.jchf.2016.03.003> PMID: 27256754
15. Nordström J, Kvernby S, Kero T, Sörensen, Harms H, Lubberink M. Left-ventricular volumes and ejection fraction from cardiac ECG-gated 15 O-water positron emission tomography compared to cardiac magnetic resonance imaging using simultaneous hybrid PET/MR. *J Nucl Cardiol* <https://doi.org/10.1007/s12350-022-03154-7>
16. Singh R. M., Singh B. M. & Mehta J. L. Role of cardiac CTA in estimating left ventricular volumes and ejection fraction. *World J Radiol* 6, 669 (2014). <https://doi.org/10.4329/wjrv.v6.i9.669> PMID: 25276310
17. Demarchi A, Neumann L, Rordorf R, Conte G, Sanzo A, Zkartal TO, et al. Long-term outcome of catheter ablation for atrial fibrillation in patients with severe left atrial enlargement and reduced left ventricular ejection fraction. <https://doi.org/10.1093/europace/euab213> PMID: 34534277
18. Hundley WG, Bluemke DA, Bogaert J, Flamm SD, Fontana M, Friedrich MG, et al. Society for Cardiovascular Magnetic Resonance (SCMR) guidelines for reporting cardiovascular magnetic resonance examinations. *Journal of Cardiovascular Magnetic Resonance* 24, 29 (2022). <https://doi.org/10.1186/s12968-021-00827-z> PMID: 35484555
19. Stassen J, Singh GK, Pio SM, Chimed S, Butcher SC, Hirasawa K, et al. Incremental value of left ventricular global longitudinal strain in moderate aortic stenosis and reduced left ventricular ejection fraction. *Int J Cardiol* 373, 101–106 (2023). <https://doi.org/10.1016/j.ijcard.2022.11.035> PMID: 36427607
20. Lih OS, Jahmunah V, San TR, Ciaccio EJ, Yamakawa T, Tanabe M, et al. Comprehensive electrocardiographic diagnosis based on deep learning. (2020) <https://doi.org/10.1016/j.artmed.2019.101789> PMID: 32143796
21. Alyounis S., Hadjileontiadis L.J., Khandoker A.H. and Stefanini C., Non-Invasive Technologies for Heart Failure, Systolic and Diastolic Dysfunction Modeling: A Scoping Review. *Frontiers in Bioengineering and Biotechnology*, Volume 11–2023, <https://doi.org/10.3389/fbioe.2023.1261022>
22. Kiranyaz S., Ince T. & Gabbouj M. Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks. *IEEE Trans Biomed Eng* 63, 664–675 (2016). <https://doi.org/10.1109/TBME.2015.2468589> PMID: 26285054
23. Rajendra Acharya U, Lih Oh S, Hagiwara Y, Tan JH, Adam M, Gertych A, et al. A deep convolutional neural network model to classify heartbeats. (2017) <https://doi.org/10.1016/j.combiomed.2017.08.022> PMID: 28869899
24. Liu W, Zhang M, Zhang Y, Liao Y, Huang Q, Chang S, et al. Real-Time Multilead Convolutional Neural Network for Myocardial Infarction Detection. *IEEE J Biomed Health Inform* 22, (2018). <https://doi.org/10.1109/JBHI.2017.2771768> PMID: 29990164
25. Andreotti F., Carr O., Pimentel M. A. F., Mahdi A. & De Vos M. Comparing feature-based classifiers and convolutional neural networks to detect arrhythmia from short segments of ECG. *Comput Cardiol (2010)* 44, 1–4 (2017).
26. Rajendra Acharya, Fujita H, Lih Oh S, Hagiwara Y, Jen, Tan H, et al. Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals. <https://doi.org/10.1007/s10489-018-1179-1>
27. Sudarshan VK, Acharya Ur, Lih Oh, Adam M, Hong Tan J, Kuang Chua C, et al. Automated diagnosis of congestive heart failure using dual tree complex wavelet transform and statistical features extracted from 2 s of ECG signals A R T I C L E I N F O. (2017) <https://doi.org/10.1016/j.combiomed.2017.01.019> PMID: 28231511
28. Gatzoulis KA, Tsiachris D, Arsenos P, Antoniou CK, Dilaveris P, Sideris S, et al. Arrhythmic risk stratification in post-myocardial infarction patients with preserved ejection fraction: the PRESERVE EF study. *Eur Heart J.* (2019) 40:2940–9. <https://doi.org/10.1093/eurheartj/ehz260> PMID: 31049557
29. Gatzoulis KA, Tsiachris D, Arsenos P, Dilaveris P, Sideris S, Simantirakis E, et al. Post myocardial infarction risk stratification for sudden cardiac death in patients with preserved ejection fraction: PRESERVE-EF study design. *Hellenic J Cardiol.* (2014) 55:361–8. PMID: 25243434
30. University of Rochester Medical Center. Telemetric and Holter ECG Warehouse (THEW). Available online at: <http://thew-project.org/databases.htm> (accessed December 24, 2019).
31. Burattini L, Burattini R. Characterization of Repolarization Alternans in the Coronary Artery Disease. In *Coronary Artery Diseases*. IntechOpen. (2012). <https://doi.org/10.5772/28766>
32. Jelinek HF, Karmakar C, Kiviniemi AM, Hautala AJ, Tulppo MP, Mäkikallio TH, et al. Temporal dynamics of the circadian heart rate following low and high volume exercise training in sedentary male

- subjects. *Eur J Appl Physiol* 115, 2069–2080 (2015). <https://doi.org/10.1007/s00421-015-3185-x> PMID: 25995100
33. Saleem S., Khandoker A. H., Alkhodari M., Hadjileontiadis L. J. & Jelinek H. F. A two-step pre-processing tool to remove Gaussian and ectopic noise for heart rate variability analysis. *Sci. Rep.* 12(1), 18396 (2022). <https://doi.org/10.1038/s41598-022-21776-2> PMID: 36319659
 34. Bazett H.C. (1920) An Analysis of the Time-Relations of Electrocardiograms. *Heart*, 7, 353–370.
 35. Christensen T. F., Randløv J., Kristensen L. E., Eldrup E., Hejlesen O. K., and Struijk J. J., ‘QT Measurement and Heart Rate Correction during Hypoglycemia: Is There a Bias?’, *Cardiol. Res. Pract.*, vol. 2010, p. e961290, Dec. 2010, <https://doi.org/10.4061/2010/961290> PMID: 21234404
 36. Liu F., Wei S., Li Y., Jiang X., Zhang Z., Zhang L., et al. ‘The Accuracy on the Common Pan-Tompkins Based QRS Detection Methods Through Low-Quality Electrocardiogram Database’, *J. Med. Imaging Health Inform.*, vol. 7, no. 5, pp. 1039–1043, Sep. 2017, <https://doi.org/10.1166/jmihi.2017.2134>
 37. Sanghavi Rohan (2023). ECG SIGNAL PQRST PEAK DETECTION TOOLBOX (<https://www.mathworks.com/matlabcentral/fileexchange/73850-ecg-signal-pqrst-peak-detection-toolbox>), MATLAB Central File Exchange. Retrieved July 28, 2023.
 38. Vapnik V. N., *Statistical Learning Theory*. Hoboken, NJ, USA: Wiley, 1998.
 39. Zhang F. & O’Donnell L. J. Chapter 7—Support vector regression. In *Machine Learning* (eds. Mechelli A. & Vieira S.) 123–140 (Academic Press, 2020). <https://doi.org/10.1016/B978-0-12-815739-8.00007-9>.
 40. Parthasarathy, S., Vaishnavi, J., Chennai, I & Princy, J. P. Predicting Heart Failure using SMOTE-ENN-XGBoost. <https://doi.org/10.1109/IDCIoT56793.2023.10053458>
 41. Davagdorj, K, Pha m, V. H, Theera-Umpon, N. & Ho Ryu, K XGBoost-Based Framework for Smoking-Induced Noncommunicable Disease Prediction. <https://doi.org/10.3390/ijerph17186513>
 42. Rasmussen C. E. & Williams C. K. I. *Gaussian processes for machine learning*. (2006). Available from: www.GaussianProcess.org/gpml
 43. McCulloch A. D. & Kerckhoffs R. C. P. Cardiac Biomechanics. *Biomedical Engineering Fundamentals* 15-1-15–30 (2014) <https://doi.org/10.1201/B15482-20/CARDIAC-BIOMECHANICS-ANDREW-MCCULLOCH-ROY-KERCKHOF>
 44. Marsden A., Feinstein J., applied C. T.-C. methods in & 2008, undefined. A computational framework for derivative-free optimization of cardiovascular geometries. *Elsevier*.
 45. Mazhari R., Omens J., . . . J. C.-C. & 2000, undefined. Structural basis of regional dysfunction in acutely ischemic myocardium. *academic.oup.com/R Mazhari, JH Omens, JW Covell, AD McCulloch Cardiovascular research, 2000*academic.oup.com*.
 46. Buhmann M. D. Regression Trees. *Encyclopedia of Machine Learning and Data Mining* 1080–1083 (2017) https://doi.org/10.1007/978-1-4899-7687-1_717
 47. Ozcan M. & Peker S. A classification and regression tree algorithm for heart disease modeling and prediction. *Healthcare Analytics* 3, 100130 (2023).
 48. Medica D. G.-B. & 2015, undefined. Lessons in biostatistics. *academia.edu D Giavarina Biochem Medica, 2015*academia.edu*.
 49. Chen T. & Guestrin C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (Association for Computing Machinery, 2016). <https://doi.org/10.1145/2939672.2939785>
 50. Hastie T., Tibshirani R. & Friedman J. Additive Models, Trees, and Related Methods. 295–336 (2009) https://doi.org/10.1007/978-0-387-84858-7_9
 51. Giavarina D. Understanding Bland Altman analysis. *Biochem Med (Zagreb)* 25, 141–151 (2015). <https://doi.org/10.11613/BM.2015.015> PMID: 26110027
 52. Alkhodari M., Jelinek H. F., Werghi N., Hadjileontiadis L. J. & Khandoker A. H. Estimating Left Ventricle Ejection Fraction Levels Using Circadian Heart Rate Variability Features and Support Vector Regression Models. *IEEE J Biomed Health Inform* 25, 746–754 (2021). <https://doi.org/10.1109/JBHI.2020.3002336> PMID: 32750938
 53. Alkhodari M, Jelinek HF, Karlas A, Soulaïdopoulos S, Arsenos P, Doundoulakis I, et al. Deep Learning Predicts Heart Failure With Preserved, Mid-Range, and Reduced Left Ventricular Ejection Fraction From Patient Clinical Profiles. *Front Cardiovasc Med* 8, 755968 (2021). <https://doi.org/10.3389/fcvm.2021.755968> PMID: 34881307
 54. Alkhodari M, Jelinek HF, Saleem S, Hadjileontiadis LJ, Khandoker AH. Revisiting Left Ventricular Ejection Fraction Levels: A Circadian Heart Rate Variability-Based Approach. *IEEE Access*. 2021; 9:130111–26.

55. Schwartz B. G., Mayeda G. S., Burstein S., Economides C. & Kloner R. A. When and why do heart attacks occur? Cardiovascular triggers and their potential role. *Hosp Pract (1995)* 38, 144–152 (2010). <https://doi.org/10.3810/hp.2010.06.308> PMID: 20890064
56. Shea, S. A, Hilton, M. F. & Muller, J. E. Day/Night Pattern of Myocardial Infarction and Sudden Cardiac Death. In *Blood Pressure Monitoring in Cardiovascular Medicine and Therapeutics* 253–291 (Humana Press). https://doi.org/10.1007/978-1-59259-978-3_11
57. Seneviratna A, Lim GH, Devi A, Carvalho LP, Chua T, Koh TH, et al. Circadian Dependence of Infarct Size and Acute Heart Failure in ST Elevation Myocardial Infarction. *PLoS One* 10, (2015). <https://doi.org/10.1371/journal.pone.0128526> PMID: 26039059
58. Ye M, Zhang JW, Liu J, Zhang M, Yao FJ, Cheng YJ. Association Between Dynamic Change of QT Interval and Long-Term Cardiovascular Outcomes: A Prospective Cohort Study. *Front Cardiovasc Med* 8, (2021). <https://doi.org/10.3389/fcvm.2021.756213> PMID: 34917661
59. Santini M. Biventricular pacing in patients with heart failure and intraventricular conduction delay: state of the art and perspectives. The European view. *Eur Heart J* 23, 682–686 (2002). <https://doi.org/10.1053/euhj.2001.2958> PMID: 11977989
60. Arsenos P, Gatzoulis KA, Laina A, Doundoulakis I, Soulaïdopoulos S, Kordalis A, et al. QT interval extracted from 30-minute short resting Holter ECG recordings predicts mortality in heart failure. *J Electrocardiol* 72, 109–114 (2022). <https://doi.org/10.1016/j.jelectrocard.2022.03.013> PMID: 35452874
61. Zemljic, G., Bunc, M. & Vrtovec, B Trimetazidine Shortens QTc Interval in Patients With Ischemic Heart Failure. <https://doi.org/10.1177/1074248409354601> 15, 31–36 (2009).
62. Shenkman HJ, Pampati V, Khandelwal AK, McKinnon J, Nori D, Kaatz S, et al. Congestive Heart Failure and QRS Duration: Establishing Prognosis Study. *Chest* 122, 528–534 (2002). <https://doi.org/10.1378/chest.122.2.528> PMID: 12171827
63. Bleeker GB, Schalij MJ, Molhoek SG, Verwey HF, Holman ER, Boersma E, et al. Relationship Between QRS Duration and Left Ventricular Dyssynchrony in Patients with End-Stage Heart Failure. *J Cardiovasc Electrophysiol* 15, 544–549 (2004). <https://doi.org/10.1046/j.1540-8167.2004.03604.x> PMID: 15149423
64. Iuliano S., Fisher S. G., Karasik P. E., Fletcher R. D. & Singh S. N. QRS duration and mortality in patients with congestive heart failure. *Am Heart J* 143, 1085–1091 (2002). <https://doi.org/10.1067/mhj.2002.122516> PMID: 12075267
65. Padeletti L, Valleggi A, Vergaro G, Lucà F, Rao CM, Perrotta L, et al. Concordant Versus Discordant Left Bundle Branch Block in Heart Failure Patients: Novel Clinical Value of an Old Electrocardiographic Diagnosis. *J Card Fail* 16, 320–326 (2010). <https://doi.org/10.1016/j.cardfail.2009.12.005> PMID: 20350699
66. Stone, P. H. ST-Segment Analysis in Ambulatory ECG (AECG or Holter) Monitoring in Patients with Coronary Artery Disease: Clinical Significance and Analytic Techniques.
67. Khot UN, Jia G, Moliterno DJ, Lincoff AM, Khot MB, Harrington RA, et al. Prognostic Importance of Physical Examination for Heart Failure in Non–ST-Elevation Acute Coronary Syndromes: The Enduring Value of Killip Classification. *JAMA* 290, 2174–2181 (2003). <https://doi.org/10.1001/jama.290.16.2174> PMID: 14570953
68. Liu Y, Wang LF, Yang XC, Lu CL, Li KB, Chen ML, et al. The long-term impact of a chronic total occlusion in a non-infarct-related artery on acute ST-segment elevation myocardial infarction after primary coronary intervention. *BMC Cardiovasc Disord* 21, (2021).