

Article

A Wearable IoT-Based Measurement System for Real-Time Cardiovascular Risk Prediction Using Heart Rate Variability

Nurdaulet Tasmurzayev^{1,*}, Bibars Amangeldy^{1,*}, Timur Imankulov¹, Baglan Imanbek¹,
Octavian Adrian Postolache² and Akzhan Konysbekova³

¹ Faculty of Information Technologies and Artificial Intelligence, Al Farabi Kazakh National University, Almaty 050040, Kazakhstan; timur.imankulov@kaznu.edu.kz (T.I.); imanbek.baglan@kaznu.kz (B.I.)

² Department of Information Science and Technology, ISCTE—Instituto Universitário de Lisboa, 1649-026 Lisbon, Portugal; octavian.adrian.postolache@iscte-iul.pt

³ JSC “Research Institute of Cardiology and Internal Diseases”, Almaty 050000, Kazakhstan; konysbekovaau@gmail.com

* Correspondence: tasmurzayev.n@gmail.com (N.T.); bibars.amangeldy@ieee.org (B.A.)

Abstract

Cardiovascular diseases (CVDs) remain the leading cause of global mortality, with ischemic heart disease (IHD) being the most prevalent and deadly subtype. The growing burden of IHD underscores the urgent need for effective early detection methods that are scalable and non-invasive. Heart Rate Variability (HRV), a non-invasive physiological marker influenced by the autonomic nervous system (ANS), has shown clinical relevance in predicting adverse cardiac events. This study presents a photoplethysmography (PPG)-based Zhurek IoT device, a custom-developed Internet of Things (IoT) device for non-invasive HRV monitoring. The platform's effectiveness was evaluated using HRV metrics from electrocardiography (ECG) and PPG signals, with machine learning (ML) models applied to the task of early IHD risk detection. ML classifiers were trained on HRV features, and the Random Forest (RF) model achieved the highest classification accuracy of 90.82%, precision of 92.11%, and recall of 91.00% when tested on real data. The model demonstrated excellent discriminative ability with an area under the ROC curve (AUC) of 0.98, reaching a sensitivity of 88% and specificity of 100% at its optimal threshold. The preliminary results suggest that data collected with the “Zhurek” IoT devices are promising for the further development of ML models for IHD risk detection. This study aimed to address the limitations of previous work, such as small datasets and a lack of validation, by utilizing real and synthetically augmented data (conditional tabular GAN (CTGAN)), as well as multi-sensor input (ECG and PPG). The findings of this pilot study can serve as a starting point for developing scalable, remote, and cost-effective screening systems. The further integration of wearable devices and intelligent algorithms is a promising direction for improving routine monitoring and advancing preventative cardiology.

Keywords: ischemic heart disease; heart rate variability; cardiovascular diseases; wearable technology; machine learning; IoT device; health monitoring; risk prediction



Academic Editor: Hisham Daoud

Received: 30 July 2025

Revised: 11 September 2025

Accepted: 28 September 2025

Published: 2 October 2025

Citation: Tasmurzayev, N.; Amangeldy, B.; Imankulov, T.; Imanbek, B.; Postolache, O.A.; Konysbekova, A. A Wearable IoT-Based Measurement System for Real-Time Cardiovascular Risk Prediction Using Heart Rate Variability. *Eng* **2025**, *6*, 259. <https://doi.org/10.3390/eng6100259>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

CVDs are the leading cause of death worldwide. According to estimates by the World Health Organization, in 2019, 17.9 million people died from CVDs, accounting for 32% of all global deaths. Of these deaths, 85% were due to heart attacks and strokes. Among the 17 million premature deaths (under the age of 70) from non-communicable diseases

recorded in 2019, 38% were caused by CVDs [1]. IHD is one of the most common forms of CVD and a major cause of mortality [2]. According to 2022 data, diseases of the circulatory system are the most prevalent among the adult population of Kazakhstan, with 3962.5 cases per 100,000 population. Of these, IHD accounts for 560.7 cases per 100,000 population. These figures confirm the high prevalence of IHD within the structure of cardiovascular pathology [3].

IHD continues to pose a significant burden on individuals and healthcare systems worldwide. The impact of this condition is considerable, contributing substantially to both mortality and morbidity [4]. Coronary artery disease (CAD), most commonly resulting from atherosclerosis, is the leading cause of IHD, which manifests as myocardial ischemia. The primary mechanism underlying IHD is obstructive atherosclerosis of the coronary arteries, leading to impaired blood flow to the heart muscle [5]. Increasingly, sleep health is recognized as a critical, modifiable risk factor for CVD. Poor sleep quality, insufficient duration, and disorders like obstructive sleep apnea directly contribute to the progression of atherosclerosis and hypertension through mechanisms involving systemic inflammation, endothelial dysfunction, and sympathetic nervous system overactivity. Given the growing global demand on healthcare systems, there is an urgent need to develop early risk stratification tools capable of identifying individuals at high risk for IHD before the occurrence of irreversible complications such as myocardial infarction or chronic heart failure (CHF) [6].

HRV is defined as the fluctuation in the duration of cardiac cycles [7]. It is a non-invasive indicator obtained through heart rhythm monitoring that provides valuable insights into the overall health status of the body [8]. HRV reflects the dynamic capacity of the heart and the general physiological ability of an individual to adapt to varying environmental conditions through compensatory mechanisms [9]. It is directly influenced by the primitive components of the ANS, particularly the parasympathetic branch, and also reflects the combined activity of both the sympathetic and parasympathetic divisions. Low HRV values have been associated with adverse cardiac events such as myocardial infarction, atherosclerosis progression, heart failure, IHD, and sudden cardiac death [10]. HRV analysis is essential for assessing the functional state of the ANS [11].

Current clinical tests used to assess coronary health are often expensive, invasive, and insufficiently effective for the timely detection of progressing coronary ischemic conditions [12]. Although analytical angiography is considered one of the most accurate procedures for identifying heart abnormalities, it is associated with high costs, potential side effects, and requires significant technological expertise. Traditional diagnostic methods are time-consuming, prone to human error, and may lead to inaccurate diagnoses, making them costly and labor-intensive [13]. HRV analysis emerges as a promising non-invasive alternative, as HRV is a recognized indicator of autonomic imbalance and a predictor of adverse cardiac events, and low HRV values have been directly linked to IHD. However, the standard diagnosis of HRV is based on 24 h Holter ECG, which limits its widespread use. Such conventional ECG systems require clinical supervision and meticulous electrode placement, which increase operating costs and inconvenience users. Against this backdrop, PPG has attracted particular attention, offering a significantly cheaper and more convenient alternative to ECG for continuous heart monitoring.

While the link between HRV and CVDs is well established, a practical method for IHD screening with consumer-grade PPG sensors is still lacking and the most informative HRV biomarkers obtainable from such devices have yet to be defined. To fill this gap, the present pilot introduces Zhurek, a fingertip PPG device designed in-house that performs on-board, real-time analysis of HRV metrics and securely streams the data to a cloud repository. Bench tests against a three-lead Holter ECG show clinically acceptable differences of -0.601 bpm for mean heart rate (HR), $+33.1$ ms for standard deviation of NN intervals

(SDNN), and -4.8 ms for root mean square of successive differences (RMSSDs). Using Zhurek, HRV recordings were collected from patients drawn from the Cardiology center in Almaty, Kazakhstan. Our findings indicate a strong link between HRV and IHD. Mutual information analysis revealed that the frequency-domain features High frequency (HF) and Low frequency (LF) have the highest statistical dependency on IHD status. To address a limited sample size, we employed a CTGAN model to generate synthetic HRV data, successfully expanding our dataset while preserving key statistical properties. We then trained several ML classifiers on our dataset. The RF model demonstrated the best performance, achieving an accuracy of 94% in distinguishing between individuals with and without IHD. A SHAP analysis confirmed the importance of frequency-domain metrics, identifying HF and LF as the most influential features for the model's predictions. These results underscore the potential of using ML with HRV analysis for the non-invasive detection of IHD. Our pilot study shows that measurements obtained with an affordable PPG device allow for the analysis of key HRV markers. While this method is not intended to replace the “gold standard”—24 h Holter monitoring—it demonstrates the potential of low-cost wearable sensors for HRV analysis. This lays the foundation for the development of future scalable and cost-effective screening and monitoring systems for IHD.

2. Literature Review

CVDs remain the leading cause of morbidity and mortality worldwide, highlighting the critical importance of early diagnosis in high-risk individuals and the development of effective preventative and interventional strategies. The diagnosis of IHD remains a complex challenge. Invasive coronary angiography is the “gold standard”; however, there is a pressing need for non-invasive, rapid, and reliable alternatives [14,15]. In this context, risk prediction methods play a key role, assessing the probability of future cardiovascular events (myocardial infarction, mortality) based on risk factor analysis, which allows for the identification of high-risk patients for timely intervention [16–18]. Recent research has focused on creating multifactorial models that integrate physiological indicators, lifestyle factors, and clinical history to improve risk assessment reliability and enable a more personalized approach [19]. ML models offer a powerful tool for these tasks, effectively integrating clinical variables, imaging data, and biomarkers to improve diagnostic and prognostic accuracy [16,20,21].

HRV, which measures the variation in time intervals between consecutive heartbeats (RR or NN intervals), is a non-invasive indicator widely used to assess cardiovascular health [22]. In the context of IHD, reductions in the time-domain indices SDNN, RMSSDs, and pNN50, along with changes in the low-frequency to high-frequency (LF/HF) ratio, correlate with myocardial injury and a higher risk of adverse events, while imbalance between LF and HF components reflects impaired autonomic regulation during ischemic episodes. This review systematizes these key HRV parameters and underscores their clinical relevance for monitoring and managing patients with IHD [6,23]. In patients with IHD and arrhythmias, HRV metrics are significantly reduced compared to healthy individuals. Notably, time-domain parameters such as SDNN, SDANN, RMSSD, pNN50, and the triangular HRV index, along with non-linear measures like α , α_1 , α_2 , SD1, SD2, Approximate Entropy (AppEn), and Sample Entropy (SampEn), show marked decreases in these patients [24]. These changes reflect impaired autonomic regulation of the heart and underline the utility of HRV analysis for evaluating cardiac function and disease progression in IHD [22].

Traditionally, the assessment of cardiovascular function relies on ECG, which records the heart's electrical activity from the skin surface using electrodes [25]. Although conventional ECG systems provide high accuracy, they require clinical supervision, careful

electrode placement, and regular calibration, which increase operating costs and reduce user convenience [25]. Over the past decade, growing demand for continuous, convenient, and low-cost solutions has stimulated the search for alternative monitoring methods [26]. Against this background, PPG has attracted particular attention because of its simple hardware implementation and easy integration into consumer devices, offering a cheaper and more convenient option for continuous monitoring in both clinical and everyday settings [27]. Breakthroughs in microelectronics and sensor technologies have led to the miniaturization of PPG sensors and their integration into wristbands, smartwatches, mobile phones, and in-ear devices, which has democratized access to continuous cardiac monitoring [28]. The additional pairing of PPG with wireless data transmission and cloud analytics provides a unique combination of affordability, portability, and convenience that conventional ECG systems cannot fully deliver [29].

Deep learning models demonstrate outstanding effectiveness in analyzing ECG signals, where they can automatically extract key features from raw data that distinguish normal from ischemic patterns. Experimental studies confirm that such models can achieve a classification accuracy exceeding 98% in distinguishing IHD and myocardial infarction from healthy states by detecting minute deviations in ST-segment morphology and QRS complex duration—key indicators of ischemia [30]. To enhance performance, hybrid architectures combining convolutional (CNN) and recurrent (RNN) layers are used, which capture both spatial and temporal dependencies in the data [31,32]. In addition to analyzing the full ECG signal, ML algorithms are widely applied to classify condition-based HRV. Among these, RF stands out for its effective handling of non-linear patterns, achieving an accuracy of 95.1% in binary classification [33]. Additionally, K-nearest neighbors (KNN) and decision trees (DTs) have shown high accuracy, up to 92.86%, in cardiovascular status assessment tasks [34–36].

Along with ECG and PPG data, visual diagnostic data such as non-contrast CT, echocardiograms, and CT angiograms are increasingly analyzed using deep learning models. These models form hierarchical representations of coronary artery anatomy, identifying subtle changes in vessel caliber and myocardial wall motion [37,38]. Architectures like autoencoders compress multidimensional data to create interpretable latent features for further analysis [38]. However, a major challenge in working with medical data is class imbalance, where there are far fewer abnormal cases than normal ones. To address this, unsupervised anomaly detection approaches are applied, such as k-means clustering, which segments data by similarity metrics and flags deviating samples as potential anomalies, thereby improving the performance of subsequent supervised classifiers [39]. Synthetic oversampling is also used: traditional methods like SMOTE balance the classes and significantly improve the performance of algorithms such as the support vector machine (SVM) [40]. Furthermore, more modern approaches based on generative adversarial networks (GANs and CTGANs) have shown even higher effectiveness compared to classic techniques, especially in SVM and logistic regression classification models [31].

Despite advancements, several challenges remain. Linear methods for HRV analysis have limited sensitivity and fail to detect complex patterns in physiological signals. Modified non-linear methods demonstrate greater effectiveness in identifying atypical patterns, including U-shaped dependencies [41]. In the domain of ML, many studies lack adequate consideration of model uncertainty, which is a critical aspect when using ML algorithms. Neglecting this factor reduces the reliability of the results and limits their reproducibility [33]. Furthermore, the relevance of some traditional diagnostic methods is being questioned; for example, in a study focused on the diagnosis of IHD, the AUC value for exercise testing was significantly lower compared to results obtained through the analysis of volatile organic compounds (VOCs). This indicates lower diagnostic relevance of

exercise testing in this context [42]. Finally, several studies have only focused on short-term outcomes. A study aimed at predicting stroke outcomes based on HRV indicators did not consider long-term cardiovascular events.

3. Materials and Methods

3.1. System Description

The proposed hybrid physiological monitoring system is engineered for continuous HRV analysis to enable the ambulatory assessment of ANS status. The system integrates a wearable sensing device with on-device signal processing and remote data logging, as shown in Figure 1.

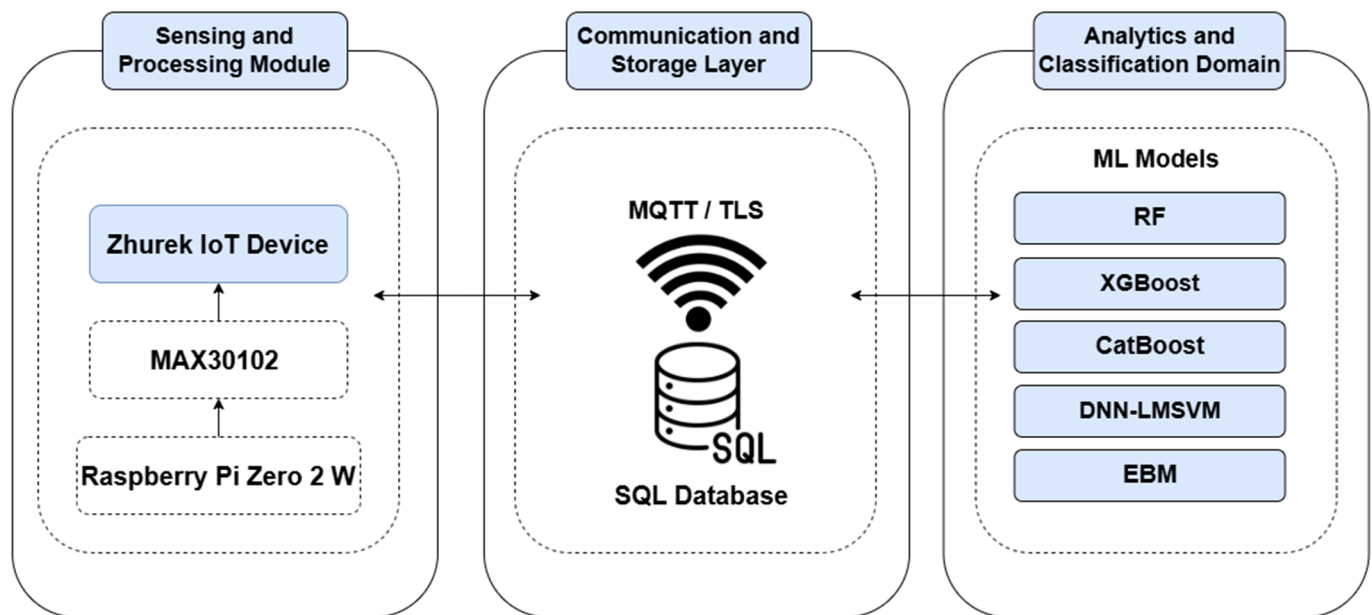


Figure 1. Architecture and description of the system.

The system is composed of three integrated subsystems: the sensing and processing module, the communication and storage layer, and the analytics and classification domain. Bidirectional arrows indicate continuous data exchange and feedback between subsystems.

At the sensing module, the Zhurek IoT device is responsible for acquiring PPG signals and computing core HRV metrics in real time. The device captures fingertip-based PPG data, processes it locally to extract time-domain features, and prepares it for wireless transmission. The embedded software processes the raw signal in real time and extracts several well-established HRV metrics, including HR, R wave to R wave intervals (RR intervals), standard deviation of normal-to-normal intervals (SDNN), and RMSSD. A detailed description of Zhurek's hardware and firmware architecture is provided in Section 3.2.

Processed HRV features are encapsulated in JavaScript Object Notation (JSON) format and transmitted via Wi-Fi using the MQTT protocol. The device publishes data to the topic zhurek/ppg/hrv, which is managed by a Mosquitto 2.0 MQTT broker hosted on a centralized server. All communication is secured using TLS 1.3 with mutual certificate-based authentication to ensure data integrity and privacy.

At the storage layer, incoming MQTT messages are parsed and stored in a relational SQL database. Each record is timestamped using a synchronized internal real-time clock, which is regularly updated via Network Time Protocol (NTP) to maintain temporal accuracy across devices. In parallel, the wearable device retains a local backup log in CSV format, providing redundancy in case of connectivity loss.

At the analytics and classification domain, the extracted HRV features are utilized to predict the risk of autonomic dysfunction associated with IHD. The stored data is periodically processed using various ML algorithms, including gradient boosting methods (XGBoost, CatBoost), RF, interpretable generalized additive models (EBMs), and hybrid architectures combining deep neural networks (DNNs) with least-mean-square support vector machines (LMSVMs). These models are trained on labeled datasets to classify patients by risk level and to detect early patterns of dysfunction. This approach enables automated preliminary diagnostics and supports wellness assessment and risk stratification in remote monitoring scenarios.

This system enables continuous monitoring and structured data analysis by integrating embedded signal processing, secure wireless transmission, and modular analytics. The use of open-source tools and commercially available components supports reproducibility and facilitates deployment in remote monitoring scenarios.

3.2. Zhurek IoT Device

Zhurek, shown in Figure 2, is a custom-engineered, non-invasive wearable device designed to capture and process PPG signals in real time. Its compact form factor, self-contained electronics, and on-device analytics make it suitable for long-term ambulatory monitoring outside clinical environments.

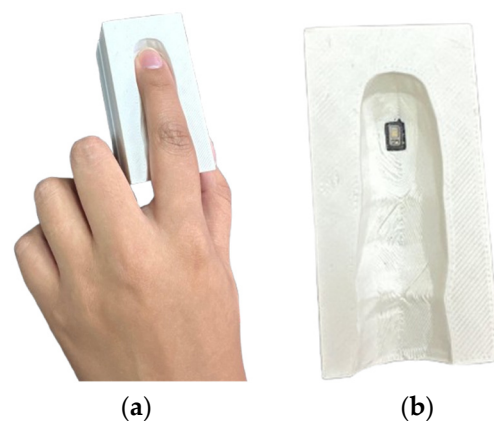


Figure 2. Zhurek IoT device: (a) The device in use, with a finger placed in the measurement slot. (b) The 3D-printed casing with the fingertip slot and optical PPG sensor exposed.

The Zhurek device integrates MAX30102 optical sensor (DFRobot Gravity: SEN0344, DFRobot, Shanghai, China) with a Raspberry Pi Zero 2 W microcontroller (ARM Cortex-A53, 1 GHz, 512 MB RAM, Raspberry Pi Ltd., Cambridge, UK) running Raspberry Pi OS Lite (64-bit, Version 6.12, Raspberry Pi Foundation, Cambridge, UK). Only the infrared channel is used for signal acquisition, sampled at 100 Hz over a hardware I²C bus (address 0×57). The sensor is enclosed in a 3D-printed PLA shell with an IR-shielded finger clip and soft elastomer padding to reduce motion artefacts and ambient interference.

The device is designed as a finger-clip wearable, intended primarily for resting-state measurements. Power can be supplied either by a rechargeable Li-Po battery (~6 h in wireless mode) or continuously through a USB power adapter, depending on the application. In terms of reliability, the system maintains >95% valid beat detection under resting conditions. Accuracy was benchmarked against a clinical-grade Holter ECG, with deviations of -0.601 bpm for HR, $+33.1$ ms for SDNN, and -4.8 ms for RMSSD, demonstrating clinically acceptable agreement.

All acquisition and processing code is implemented in Python 3.11, with smbus2 used for I²C communication. The raw PPG signal undergoes baseline correction and smoothing (via moving average filtering). A derivative-based peak detection algorithm derived from

HeartPy identifies cardiac cycles, with physiological validation applied to exclude outliers. RR intervals are calculated from peak timestamps, and time-domain HRV metrics—HR, SDNN, and RMSSD—are computed in 30 s windows with a 5 s step. Frequency-domain metrics (LF, HF, LF/HF) and Max_HR are obtained offline together with the anthropometric features of BMI and age, forming an eight-item feature vector.

Each result is serialized as JSON object and published via MQTT. In addition to live data transmission, the device logs the results locally in CSV format as a fallback mechanism. All data points are timestamped using a real-time clock (RTC) synchronized periodically via NTP.

Zhurek delivers optimal signal quality and high physiological accuracy under resting conditions. Resting monitoring reduces motion artefacts and yields stable autonomic patterns, ensuring reliable HRV computation. Validation results align with published evidence showing that resting acquisition provides the greatest accuracy and reproducibility for HRV, gas exchange, and metabolic rate measurements [43–45]. Remote HR and HRV recorded in this state correlate closely with ECG readings, and baseline metabolic rate together with respiratory exchange ratio remains stable and accurate during steady rest [45]. Physiological data collected under these conditions faithfully reflect ANS activity and serve as a robust baseline for IHD risk surveillance.

The clinical utility and predictive value of numerous HRV parameters are well established in the existing literature. Prior research [45], primarily using the gold-standard ECG, has confirmed that metrics such as SDNN, RMSSD, and the triangular index are significantly associated with patient outcomes. The primary challenge, however, lies in translating this diagnostic power from clinical-grade ECGs to convenient, non-invasive wearable devices.

To assess the suitability of the Zhurek device for HRV analysis, a validation study was conducted by comparing its readings against a reference clinical-grade three-lead Holter monitor. The devices demonstrated a high degree of concordance, with the signal trends from Zhurek closely tracking the Holter ECG, as shown in Figures 3 and 4.

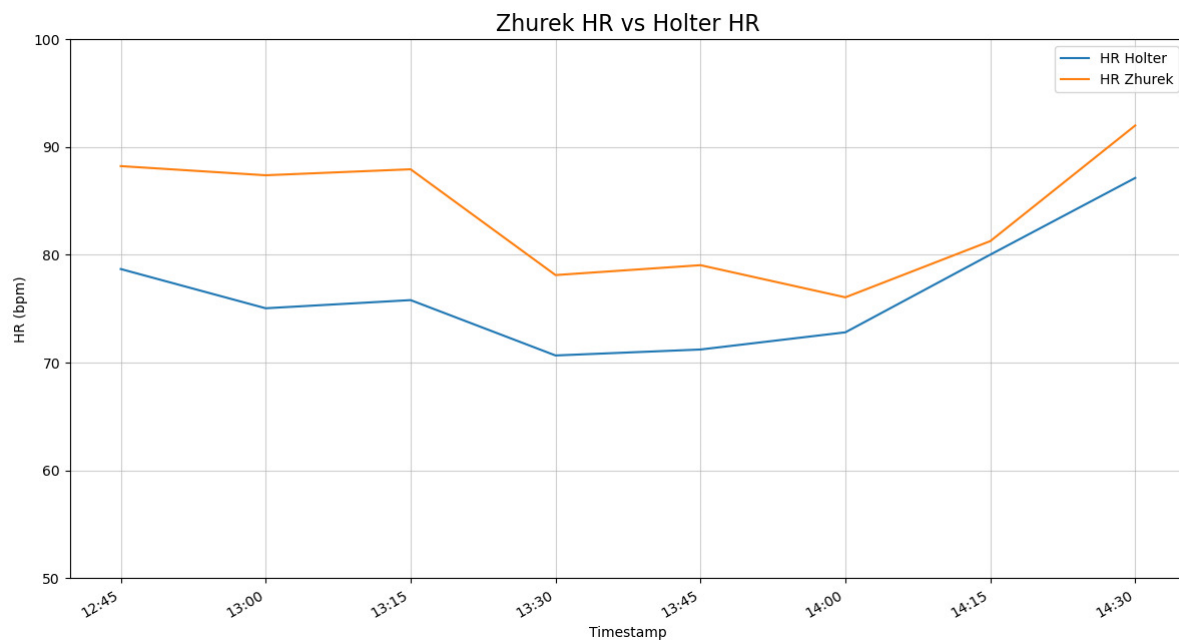


Figure 3. Time-series comparison of HR between the Zhurek device and the Holter monitor.

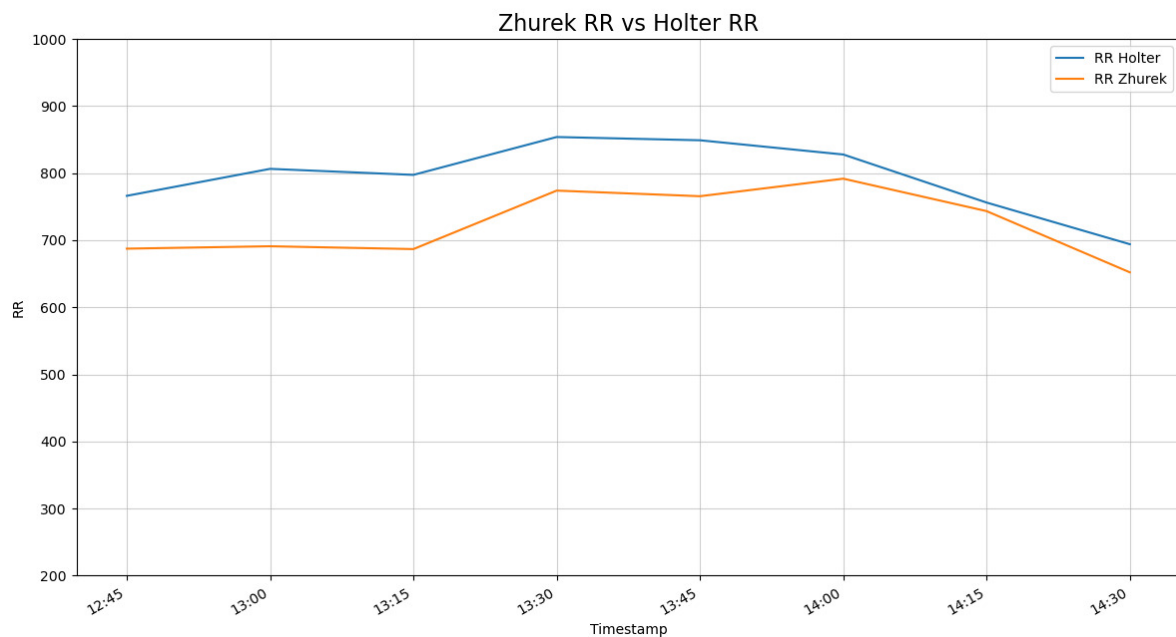


Figure 4. Time-series comparison of RR intervals between the Zhurek device and the Holter monitor.

The data reveals a strong temporal correlation, confirming that the Zhurek device accurately captures the dynamic fluctuations of cardiac rhythm. A minor, stable offset in the absolute values was observed, which characterizes the inherent difference between the PPG and ECG measurement techniques. Given that the primary goal of HRV analysis is to assess the variability in rhythm rather than absolute values, this high degree of trend alignment validates the Zhurek device as a reliable tool for its intended application.

These findings are consistent with a growing body of research focused on validating PPG-based sensors. Authors of a study [46] also validated a wearable device, demonstrating that its HRV readings closely align with reference ECG data and can serve as a valid substitute for longer, standard measurements. Further reinforcing this, another study [47] found that the accuracy of certain PPG-derived HRV parameters could be adequate for patient monitoring, underscoring the importance of parameter-specific evaluation.

3.3. Study Population and Data Collection

To ensure the robust training and evaluation of ML models for IHD prediction, HRV data were collected from two independent cohorts: a clinical group of patients with confirmed CVDs (300 patients) and a healthy control group. This dual-cohort design allows the models to learn patterns of autonomic dysfunction observed in real-world pathological states while distinguishing them from the normal physiological variability found in healthy individuals.

Monitoring of both cohorts was conducted using high-quality RR-interval acquisition devices to ensure the consistency and reliability of HRV measurements. Recordings from participants diagnosed with IHD and from healthy volunteers were collected under controlled laboratory conditions using a BTL-08 Holter ECG system (BTL Industries, Stevenage, Hertfordshire, UK) and the Zhurek IoT device (Almaty, Kazakhstan). The BTL-08 is a multi-lead ambulatory ECG recorder with up to 12-lead configuration, continuous monitoring for 24–48 h, and a sampling frequency of 1000 Hz at 12-bit resolution. It is CE-certified for diagnostic use and has validated performance in arrhythmia and ischemia detection, making it a widely adopted gold-standard reference for HRV research.

HRV data of adult inpatients with confirmed cardiovascular conditions were acquired at the Research Institute of Cardiology and Internal Diseases (Almaty, Kazakhstan). Di-

agnoses were established according to clinical protocols under the supervision of the institute's cardiology department. Continuous recordings were collected using the BTL-08 Holter ECG monitors and the Zhurek device.

Participants represented both early and advanced stages of cardiovascular pathology. This broad distribution increases population heterogeneity and supports the development of generalizable ML models. All recordings were stored as high-resolution numerical RR-interval files, and the resulting dataset already includes the key HRV variables—HR, RR intervals, SDNN, and RMSSD—automatically computed and ready for downstream analysis.

To establish a physiological baseline for HRV under normal autonomic conditions, data were collected from healthy volunteers. All participants reported no history of cardiovascular, neurological, or metabolic disorders. To minimize confounding factors, participants were instructed to abstain from alcohol, tobacco, caffeine, and intense physical activity for at least 24 h prior to data collection, and to maintain regular sleep (7–8 h) the night before. Participants were excluded if they had an acute illness, failed to meet the preparation criteria, or if signal recordings showed excessive artifacts.

Descriptive statistics and categorical characteristics of the healthy control group are summarized in Tables 1 and 2.

Table 1. Descriptive statistics of participant characteristics.

Characteristics	Mean \pm Standard Deviation
Height (cm)	173.8 \pm 10.41
Weight (kg)	71.8 \pm 18.65
BMI	23.65 \pm 7.94

Table 2. Categorical characteristics distribution of participants.

Characteristics	Categories
Gender	Female: 9, Male: 11
Genetic marker	None
Harmful Habits	Yes: 4, No: 16

3.4. Data Preprocessing

In this pilot study, data preprocessing involved the integration of two distinct cohorts: a healthy control group and a group of patients diagnosed with IHD. Large Language Models (LLMs) were used to process the raw patient data, enabling the transformation and structuring of the initial information to extract the necessary parameters for analysis. The dataset includes important HRV features such as SDNN, percentage of successive normal-to-normal intervals that differ by more than 50 milliseconds (pNN50), and RMSSD, along with frequency-domain measures including LF, HF, and the LF/HF ratio. To balance the dataset and improve model robustness, data from healthy participants were augmented using a conditional tabular generative adversarial network (CTGAN), increasing the healthy group records from 20 to 200. As a result, the total dataset consisted of 500 observations.

CTGAN is a generative model specifically designed to synthesize realistic tabular data, including both continuous and categorical features, by conditioning the generation process on selected discrete values. The model utilizes a conditional generator to resample imbalanced columns and employs mode-specific normalization to better model multi-modal distributions. The objective function of CTGAN follows the standard GAN minimax formulation, conditioned on discrete variables:

$$\min_{\mathbf{D}} \max_{\mathbf{G}} \mathbb{E}_{\mathbf{x} \sim \mathbf{P}_{\text{data}}} [\log \mathbf{D}(\mathbf{x} | \mathbf{c})] + \mathbb{E}_{\mathbf{z} \sim \mathbf{P}_{\mathbf{z}}} [\log (1 - \mathbf{D}(\mathbf{G}(\mathbf{z} | \mathbf{c})))] \quad (1)$$

where G is the generator; D is the discriminator; c is the conditional vector (discrete feature); and z is the noise input [48].

CTGAN has repeatedly demonstrated its effectiveness as a data augmentation method. In a recent medical-data study, CTGAN presented a better performance (AUC, F1-score, precision, etc.) than ROS, SMOTE, and ADASYN, outperforming 17 baseline models [31]. It has been used to generate samples of the minority class, significantly improving the performance of various ML models [49]. Experimental results showed that the use of CTGAN led to substantial improvements in classification quality based on precision, recall, and F1-score metrics compared to traditional data augmentation methods. This makes the model an optimal choice for enhancing reliability in situations with limited original data [14].

Prior to model training, the data were preprocessed by removing all rows with missing values. This approach eliminates incomplete records, reducing the risk of bias caused by inaccurate or partial information. Removing rows with missing data ensures more reliable model training, as only complete and valid observations are retained—an especially important consideration in the context of medical data analysis.

To assess the outcome of class balancing, a class distribution plot was created after applying CTGAN, showing the number of records classified as having IHD and those without it (No IHD). Figure 5 demonstrated improved class balance, which contributed to enhanced performance and generalizability of the ML models.

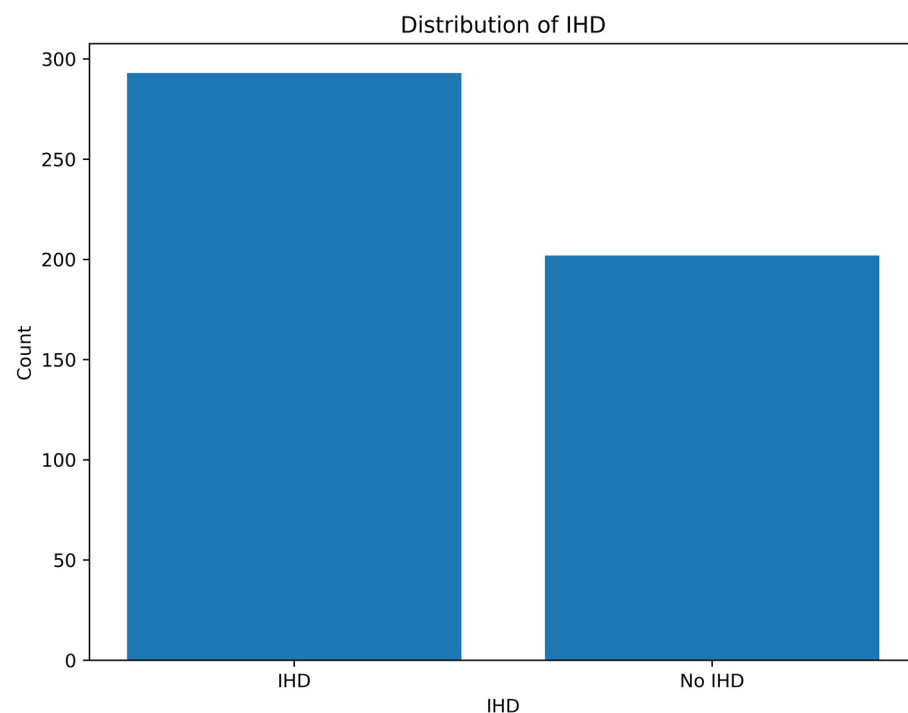


Figure 5. Distribution of IHD.

3.5. Feature Analysis and Selection

For the analysis of cardiological data, several ML models were utilized, including RF, a DNN-LMSVM, XGBoost, CatBoost, and EBMs. These models were selected based on their proven effectiveness in handling complex, non-linear relationships in biomedical data, as well as their capability to process structured and high-dimensional features. DNNs uncover hidden interactions and complex dynamics in systems purely from observational data, allowing analysis without relying on predefined models [50]. XGBoost and RF utilize ensemble learning to enhance the accuracy and reliability of assessment [15]. Moreover,

EBMs are applied to create interpretable models through generalized additive modeling, facilitating the understanding of feature contributions.

The classification task was formulated as a multi-class problem aimed at categorizing patients into different cardiovascular risk groups based on clinical and physiological features.

To ensure optimal hyperparameter settings and model robustness, grid search combined with cross-validation ($CV = 5$) was performed. This approach systematically explores various hyperparameter combinations and evaluates model generalizability across different data splits, reducing overfitting and improving predictive accuracy.

The DNN-LMSVM model leverages deep hierarchical feature extraction combined with the robust margin maximization of LMSVMs to capture subtle patterns in the data. Ensemble methods such as RF, XGBoost, and CatBoost were incorporated to improve predictive performance and reduce overfitting by aggregating multiple decision trees. Additionally, EBMs were applied to build interpretable models with the ability to analyze feature contributions through generalized additive modeling. To enhance interpretability, SHapley Additive exPlanations (SHAP) values were computed for tree-based models, allowing quantitative assessment of feature impact on individual predictions.

Data preprocessing included encoding categorical variables to ensure compatibility across all models. Target labels representing cardiovascular risk categories were encoded into numerical classes to facilitate supervised learning. The dataset comprised numerous clinical indicators and physiological measurements relevant to cardiovascular health.

Model development and training were conducted in Python using libraries including Scikit-learn (for RF and EBMs), XGBoost, CatBoost, and PyTorch (for DNN-LMSVM). The performance of the models was assessed using four primary evaluation metrics. Accuracy quantifies the overall proportion of correct predictions made by the model. Precision measures the proportion of true positive instances among all instances predicted as positive. Recall evaluates the model's ability to identify all actual positive cases within the dataset. The F1-score, as the harmonic mean of precision and recall, provides a balanced measure that captures both the model's precision and sensitivity, thereby enabling a comprehensive assessment of its reliability and predictive effectiveness.

4. Results

4.1. Participant Characteristics and HRV Data Collection

Descriptive statistics of key HRV metrics for both the healthy control group and IHD group are summarized in Table 3, respectively. The analyzed features include the time-domain measures SDNN, PNN50, and RMSSD, along with the frequency-domain measures LF, HF, and the LF/HF ratio. These values provide an overview of the ANS activity in both populations and highlight potential differences in HRV patterns between healthy individuals and those with IHD [51].

Table 3. Summary of HRV metrics in IHD and healthy control group.

HRV Features	Mean \pm Standard Deviation of IHD Group	Mean \pm Standard Deviation of Healthy Group
SDNN	64.55 \pm 26.43	61.25 \pm 23.14
PNN50	4.82 \pm 5.37	30.55 \pm 21.13
RMSSD	25.22 \pm 12.92	61.44 \pm 33.06
LF	0.16 \pm 0.08	0.03 \pm 0.03
HF	0.16 \pm 0.08	0.03 \pm 0.03
LF/HF	1.06 \pm 0.33	0.96 \pm 0.43

As shown in Figure 6, the mutual information (MI) analysis was performed to evaluate the statistical dependency between HRV features and the presence of IHD. Among the

analyzed features, the frequency domain metrics LF and HF components demonstrated the highest MI scores of 0.34 and 0.31, respectively, indicating a stronger association with IHD status. The PNN50 ratio showed a moderate MI value of 0.19, suggesting it carries some dependency with the target but less than its individual components. In contrast, time domain features such as SDNN (0.06), RMSSD (0.059), and LF/HF (0.093) exhibited lower MI values, reflecting weaker dependency.

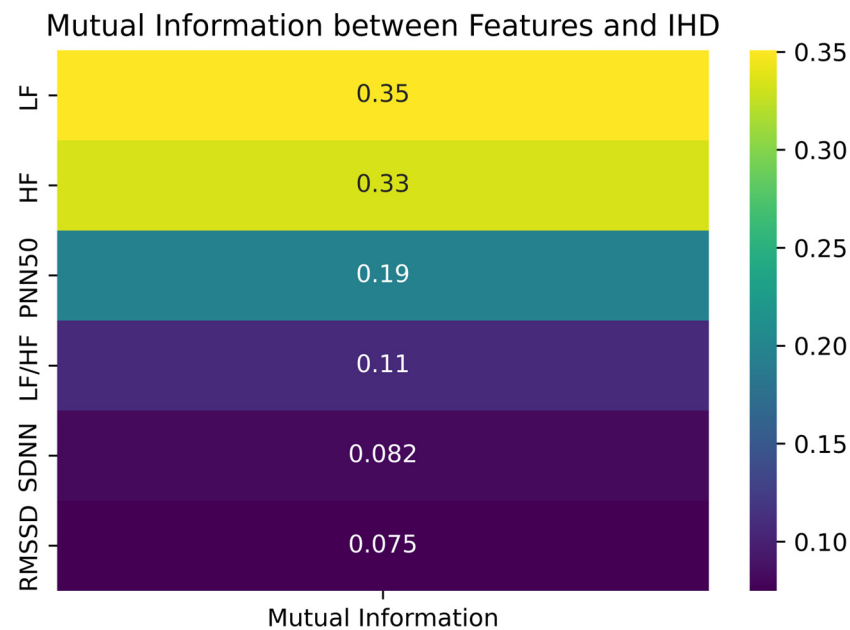


Figure 6. Mutual information scores between HRV features and IHD.

4.2. Data Augmentation and Comparison

To overcome the limited sample size of the healthy control group, the CTGAN model was employed to generate synthetic HRV data. The original dataset, comprising 50 healthy individuals, was expanded to 250 samples using this generative approach. The CTGAN model was trained for 2000 epochs to ensure adequate learning of the underlying data distribution. This augmentation allowed for more balanced model training. It is important to note that the generated data were used only for training the model, while exclusively real data were used for testing. As shown in Figure 7, the generated data closely follow the original distributions, indicating that the synthetic data successfully captures the general structure and range of the original measurements. Most features, such as SDNN, RMSSD, and LF, show strong alignment between the two datasets. However, slight differences are noticeable in features like PNN50 and LF/HF, where the generated data appear more peaked or slightly shifted. It is important to note that synthetic data is not expected to replicate the original dataset with complete precision. In fact, introducing some degree of dissimilarity is both acceptable and desirable, particularly in the context of health data where protecting patient privacy and preventing identity disclosure are critical [16]. Despite minor variations, it is essential that the synthetic dataset preserves the core statistical properties and inter-feature dependencies of the real data. In this case, the synthetic HRV metrics retained the necessary structure and distribution to support meaningful clinical insights and reliable model development.

Table 4 presents a comparison between the original and synthetic datasets based on key physiological features. For each feature, mean values, variances, and p -values from the Student's t -test are provided to assess whether there are significant differences between the two datasets.

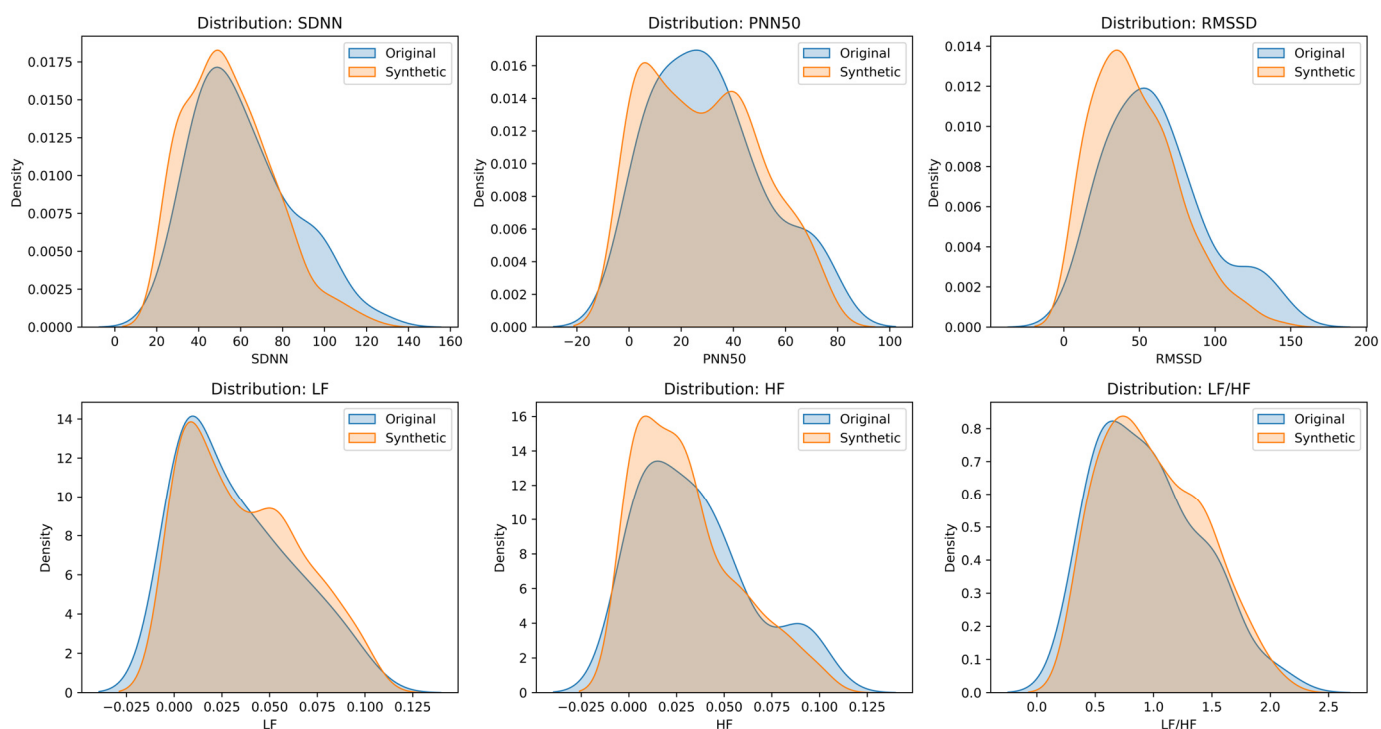


Figure 7. Distribution of key HRV metrics (SDNN, RMSSD, PNN50, LF, HF, LF/HF) in original vs. synthetic data.

Table 4. Comparison of statistical properties between original and synthetic datasets for physiological features.

Feature	Original Mean	Synthetic Mean	Mean Difference	Original Variance	Synthetic Variance	Variance Difference	<i>p</i> -Value
SDNN	61.2493	54.7188	6.530506	535.3021	435.9258	99.37634	0.052138
PNN50	30.5490	28.6446	1.904339	446.304	460.7659	−14.4619	0.572923
RMSSD	61.4390	47.7181	13.7209	1092.974	785.4974	307.477	0.002798
LF	0.032	0.0363	−0.00423	0.000829	0.00086	-3.1×10^{-5}	0.359098
HF	0.034	0.0299	0.004487	0.00083	0.000693	0.000137	0.287626
LF/HF	0.9572	1.0007	−0.04346	0.189792	0.183772	0.00602	0.521613
Max_HR	93.6760	91.8285	−1.8475	158.1155	89.1708	−68.9448	0.5295
BMI	23.6450	22.8590	−0.7860	32.4784	33.0017	0.5233	0.2892

For most features (SDNN, PNN50, LF, HF, LF/HF), the *p*-values are above 0.05, indicating no statistically significant differences between the original and synthetic data. However, for the RMSSD feature, the *p*-value is 0.002798, which indicates a statistically significant difference in the distributions of this metric.

Thus, the synthetic dataset preserves the key statistical properties of most of the original features. It can be used for further research; however, the statistically significant difference in the RMSSD metric should be taken into account, as it may affect analyses related to this specific feature.

4.3. Machine Learning Model Classification Performance

To enhance class balance and improve the robustness of model training, the synthetic data samples generated via CTGAN were merged with the original dataset. This augmented dataset was used to train a set of five ML classifiers: RF, CatBoost, XGBoost, DNN-LMSVMs, and Explainable Boosting Machine (EBM). The aim of this classification task was to use HRV features to determine whether IHD is present (1) or absent (0). Following model training, performance evaluation was carried out using 5-fold cross-validation.

As illustrated in Table 5, four standard evaluation metrics were used to assess classification quality: accuracy, F1-score, precision, and recall. All models demonstrated solid performance in distinguishing between individuals with and without IHD. Among the tested classifiers, the RF was the best-performing model, achieving an accuracy of 90.82%, with a corresponding precision of 92.11%, recall of 91.00%, and an F1-score of 90.11%. The CatBoost, XGBoost, and EBM models showed similar results with an accuracy of 88.78%. The DNN-LMSVM classifier achieved an accuracy of 84.69%.

Table 5. Accuracy, precision, recall, and F1-score values obtained for ML classifiers in the prediction of IHD based on HRV metrics.

Evaluation Metrics	RF	CatBoost	XGboost	DNN-LMSVM	EBM
Accuracy	90.82%	88.78%	88.78%	84.69%	88.78%
Precision	92.11%	90.01%	90.68%	86.00%	88.78%
Recall	91.00%	88.95%	89.00%	84.31%	88.78%
F1-Score	90.11%	87.91%	87.64%	85.15%	87.91%

The confusion matrix for the model is presented in Figure 8, illustrating the distribution of true positives, true negatives, false positives, and false negatives for the binary prediction task. Out of 48 actual class 0 instances, all 48 were correctly classified as class 0 (true negatives), with zero false positives. For class 1, the model correctly predicted 41 out of 50 instances (true positives), with 9 instances misclassified as class 0 (false negatives).

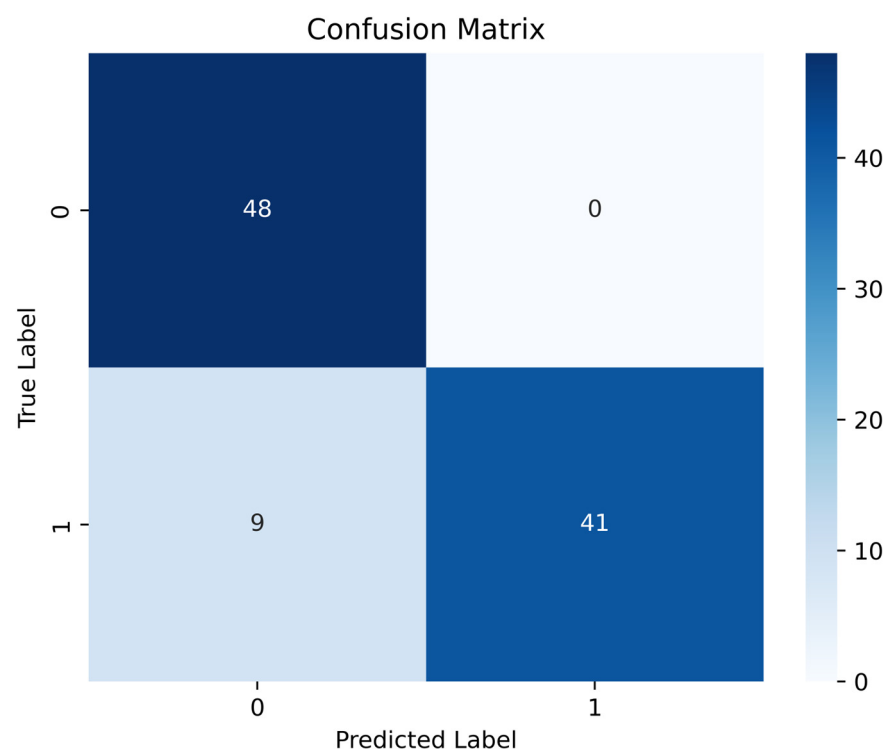


Figure 8. Confusion matrix of random forest for IHD classification.

The evaluation metrics shown in the ROC and precision–recall curves, as indicated in Figure 9, demonstrate strong overall model performance. The ROC curve shows a high true positive rate across most thresholds, with an AUC of 0.98, reflecting excellent discriminative ability between the two classes. Using Youden’s index [52], the optimal decision threshold for IHD risk prediction was identified as 0.422. At this cutoff, the model achieved a sensitivity of 88% and a specificity of 100%. This threshold can be used to stratify individuals into

“high-risk” (probability > 0.422) and “low-risk” (probability ≤ 0.422) groups, providing clinically interpretable guidance for screening and early intervention. The precision–recall curve demonstrates near-perfect performance, with precision remaining at 1.0 across the majority of recall values before only slightly declining at very high recall levels. This indicates the model is highly effective at minimizing false positives while still achieving strong sensitivity. These results indicate a high level of overall accuracy, with strong sensitivity and specificity.

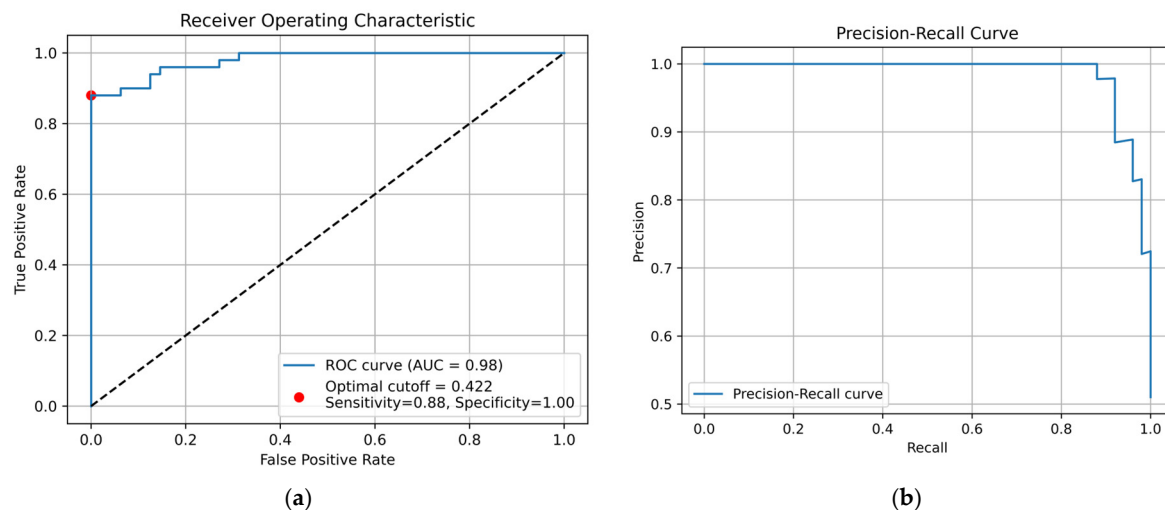


Figure 9. Model evaluation metrics: (a) ROC curve demonstrating the trade-off between true positive and false positive rates; (b) precision–recall curve illustrating the balance between precision and recall across thresholds.

Figure 10 presents the decision boundaries and feature separability of the classification model using PCA and t-SNE projections, respectively.

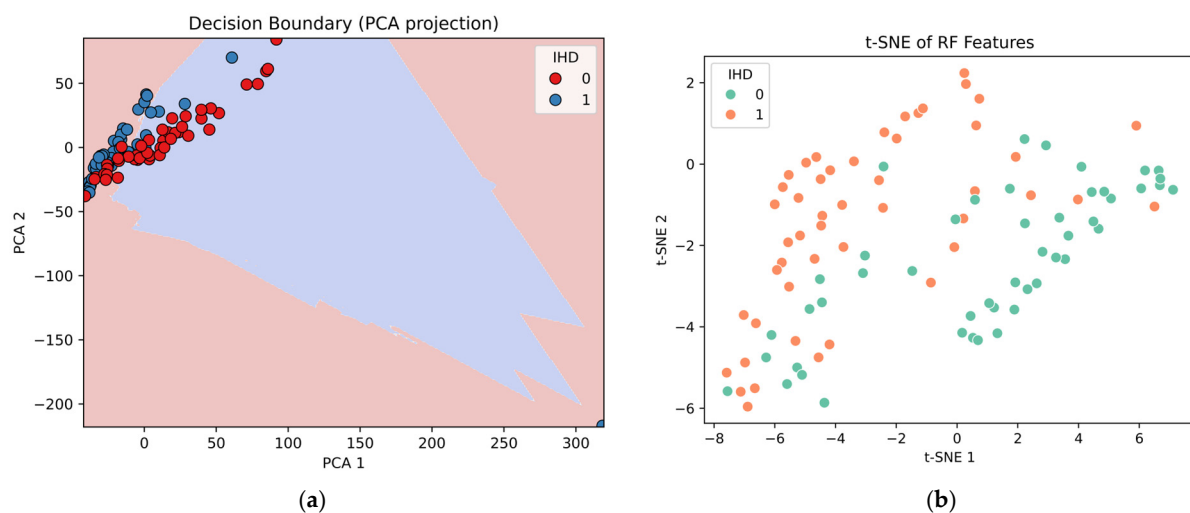


Figure 10. Feature space visualization: (a) PCA projection showing the model’s decision boundary and class distribution; (b) t-SNE projection illustrating the separability of classes based on learned feature representations.

In the PCA-based visualization, the background shading indicates the regions where the model predicts each class, while the data points are plotted according to their true labels (red for class 0, blue for class 1). In the t-SNE plot, data instances are projected into a non-linear two-dimensional space to better visualize the data structure learned by the RF model.

Both visualizations show that classes 0 and 1 generally form distinct clusters; however, there is a significant overlap between them. This overlap corresponds to the regions where the model faces higher uncertainty in classification, which is consistent with the results obtained from the confusion matrix.

The histogram of decision scores shown in Figure 11 illustrates the model's ability to distinguish between the negative and positive classes based on predicted probabilities. The distribution for the negative class (in red) is sharply concentrated near zero, while the distribution for the positive class (in green) is primarily distributed between 0.7 and 0.9. This clear separation between the two distributions indicates strong class discrimination, despite a small overlap. Such a distribution suggests the model produces confident predictions and maintains high reliability in distinguishing between the two classes.

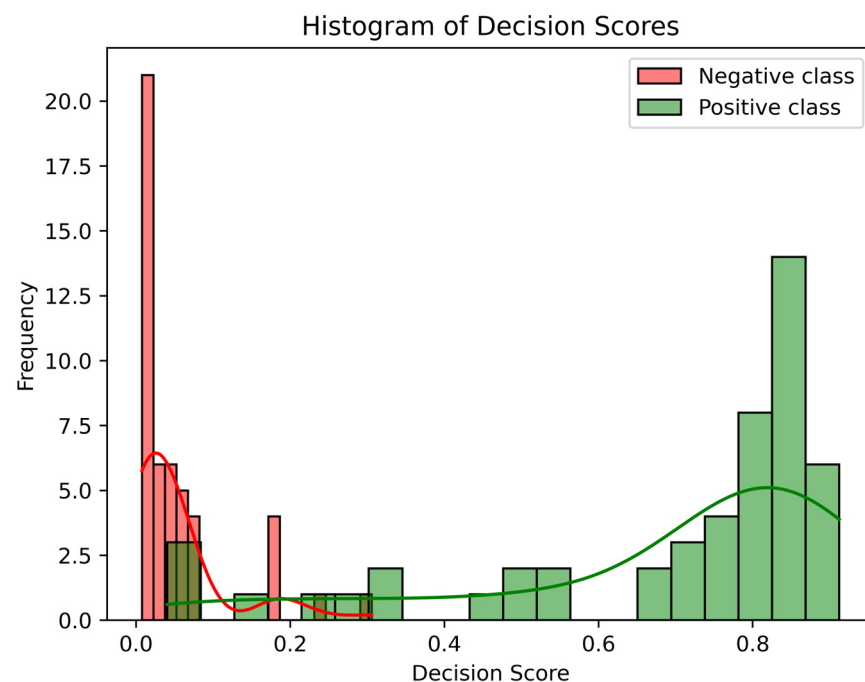


Figure 11. Histogram of decision scores: class-wise distribution of decision scores for model predictions.

To evaluate the influence of individual HRV features on IHD prediction, SHAP was applied to the trained classification model. The SHAP summary plot in Figure 12 presents the mean absolute SHAP values, highlighting the relative contribution of each feature to the model's output. HF exhibited the highest SHAP value, making it the most influential feature in the prediction task. LF ranked second in importance. They were followed, in descending order of contribution, by PNN50, SDNN, and RMSSD. The LF/HF ratio was found to have the lowest SHAP value, indicating the least impact on the model's decision-making process.

Figure 13 presents the SHAP dependence plots for six HRV features. They show how the value of each feature (X-axis) influences the model's prediction (Y-axis), with the color of the points reflecting the level of another important feature—HF (high-frequency power), which is an indicator of parasympathetic activity. Analysis of the plots reveals complex interactions, showing that the influence of nearly every metric on the prediction is highly dependent on the background parasympathetic activity. For instance, the plots for LF and LF/HF show that low HF levels (blue dots) are associated with negative SHAP values, which increases the likelihood of an IHD prediction. This suggests that the model considers an imbalance towards the sympathetic system as a risk factor, especially with weak parasympathetic activity. Conversely, for SDNN, PNN50, and RMSSD, an opposite interaction trend is observed: high HF levels (pink and red dots) are linked to higher

(more positive) SHAP values, which decreases the likelihood of an IHD prediction. This indicates that the model considers these metrics as signs of health, particularly when they are supported by strong overall parasympathetic activity. These interactions demonstrate that the model has learned to evaluate not just individual values, but their combinations and the overall autonomic balance, making the prediction more accurate.

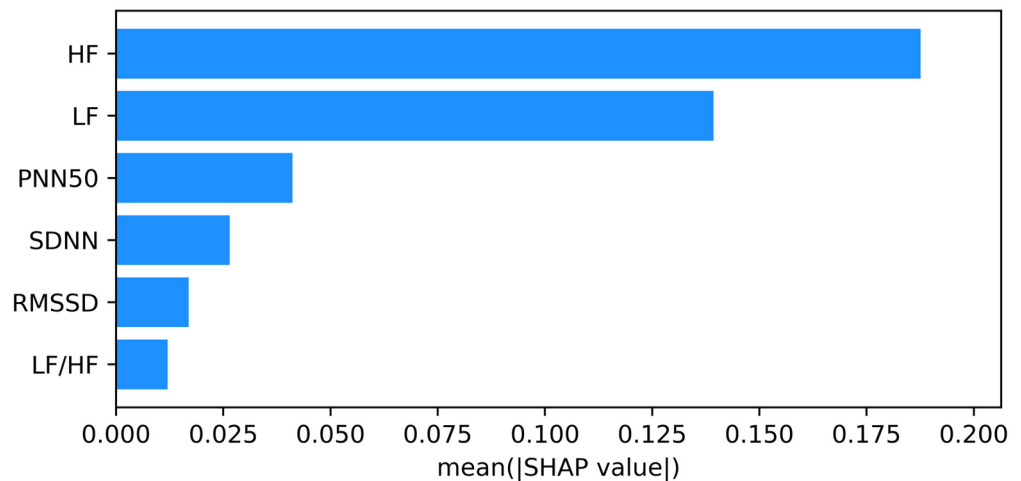


Figure 12. Feature importance analysis using SHAP.

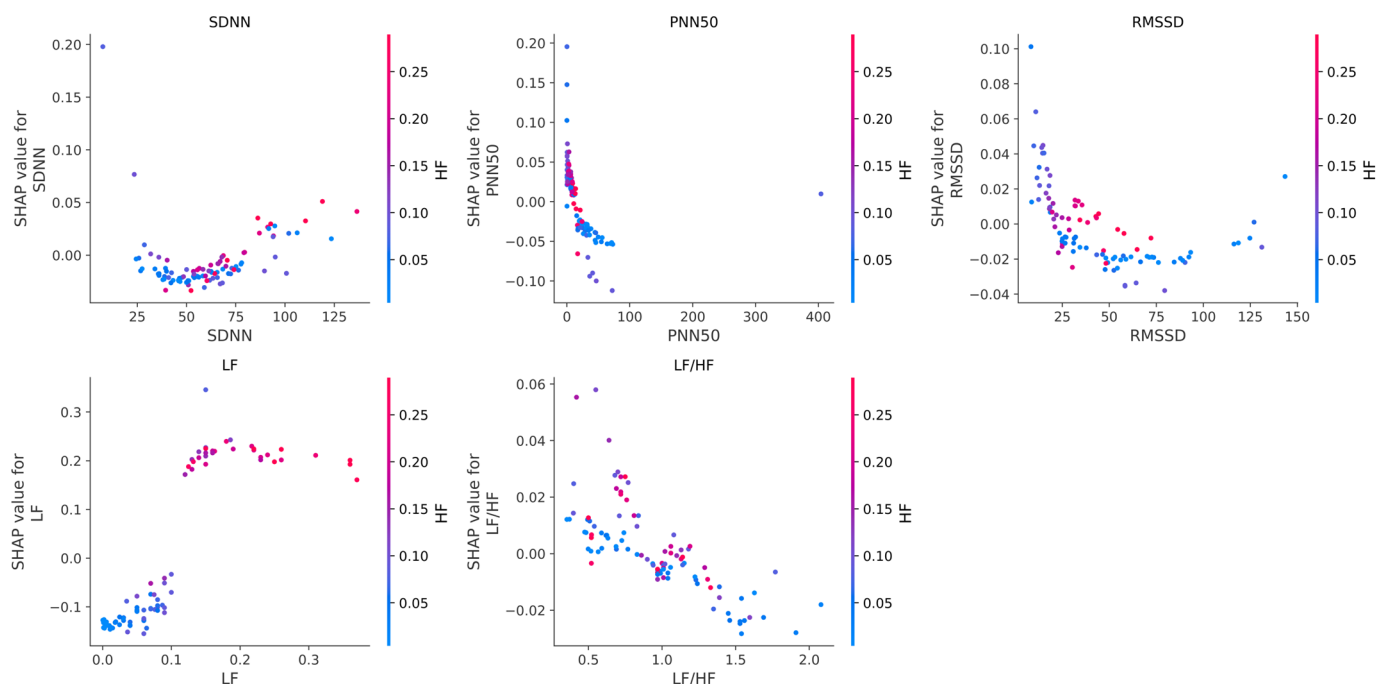


Figure 13. SHAP dependence plots for HRV features colored by HF values.

5. Discussion

A central innovation of this study is the development and rigorous clinical validation of the Zhurek IoT device, a custom-engineered, non-invasive tool designed to overcome the practical limitations of conventional cardiovascular monitoring. While the 24 h Holter ECG is the gold standard [53] for HRV analysis, its application is often restricted by high costs [54], user inconvenience [55], and the need for clinical supervision. The Zhurek device was specifically engineered to bridge this gap, offering a low-cost, ambulatory solution that captures high-fidelity PPG signals and performs real-time, on-device processing of critical HRV metrics. Thus, a primary objective was to evaluate the device's performance

and reliability under real-world conditions, establishing its viability as an accessible tool for scalable screening. To empirically prove its clinical utility, the Zhurek device was benchmarked directly against a gold-standard three-lead Holter ECG monitor in a supervised clinical setting. Concurrent HRV data were collected from both patients diagnosed with IHD and healthy controls, enabling a direct, head-to-head comparison. The results demonstrated a high degree of concordance, with signal trends from the Zhurek device closely tracking those recorded by the Holter ECG. This strong temporal correlation confirms that our device accurately captures the dynamic fluctuations in cardiac rhythm essential for meaningful HRV analysis.

The ML classification results further demonstrate that a custom, low-cost PPG system can classify IHD with high performance (AUC = 0.98) and strong signal fidelity. Specifically, it achieved a low absolute deviation of -4.8 ms for RMSSD and $+33.1$ ms for SDNN against a Holter ECG. This performance is particularly noteworthy when compared to prior validations, such as the 24 h wrist-recording study in [47], which reported significant relative mean absolute errors for high-frequency measures like RMSSD ($\approx 17\%$) and the LF/HF ratio ($\approx 36\%$). Our system's minimal deviation for RMSSD suggests a robust capability in capturing sensitive beat-to-beat variations, a known challenge for PPG technology.

Furthermore, while other foundational studies have focused on the technical nuances of signal fidelity—such as achieving a low RR-series error standard deviation (≈ 5.4 ms) with smartphone PPGs [56] or highlighting the necessity of up-sampling to 200 Hz to improve RMSSD agreement [46]—their primary objective remained signal-level validation. Our research builds upon these findings by applying the derived HRV metrics to a downstream clinical classification task. Consequently, our approach is distinct in its combination of a custom, low-cost hardware platform, validated high performance for sensitive HRV metrics, and its successful deployment in a ML model for the specific clinical purpose of IHD risk stratification.

This successful validation shows that an affordable, user-friendly PPG sensor can reliably detect the physiological differences between IHD and healthy states, establishing the Zhurek device as a foundational technology for future remote and cost-effective screening systems.

The mutual information analysis revealed that frequency-domain features, particularly LF and HF, were more strongly associated with IHD. To overcome the imbalance in class distribution, particularly due to the small number of healthy control participants, we employed CTGAN to generate synthetic HRV data. While synthetic data cannot fully replicate the complexity of physiological signals, it can approximate their statistical structure and support more balanced learning during model development.

We compared the results of CTGAN-ENN with the conventional hybrid techniques SMOTE-ENN and ADASYN-ENN, and CTGAN-ENN consistently delivered superior AUC, F1-score, and G-mean across six customer-churn datasets [57]. Our experiments further indicated that classical generative algorithms like SMOTE and ADASYN only worked well in a single dataset, whereas CTGAN-based methods remained robust across all scenarios [57].

Moreover, the use of synthetic samples provides a practical solution to privacy concerns, reducing the risk of re-identification in health datasets, thereby supporting ethical data sharing practices. Additionally, it facilitates broader access to healthcare datasets, enabling researchers to conduct timely investigations even when access to real clinical data is restricted. Synthetic datasets also serve as a valuable resource for developing and testing software solutions, especially in scenarios where realistic data are unavailable or insufficient. Despite these benefits, synthetic data must be applied with careful consideration of its limitations. Its fidelity depends heavily on the underlying generative model

and the quality of the original data used for training. Not all synthetic samples fully replicate the distributions or dependencies present in real datasets, which may introduce biases or affect downstream model performance. Moreover, there remains a recognized risk of data leakage if synthetic generation is not properly managed, and overreliance on imputation models may obscure true physiological variability [16–18]. In this study, slight discrepancies were observed in features such as PNN50 and LF/HF, underscoring the need for cautious interpretation.

In probability and information theory, mutual information measures how much two random variables depend on each other. It specifically quantifies how much information is gained about one variable by knowing the value of the other [20]. Applied to our dataset, this analysis demonstrated that frequency-domain HRV features (LF and HF) exhibit the strongest dependency on the target variable, suggesting they carry the most relevant physiological signals associated with autonomic imbalance in IHD. Evaluation metrics, including the ROC and precision–recall curves, confirmed the reliability of the classifiers. The high area under the ROC curve ($AUC = 0.98$) reflects strong sensitivity and specificity, while the histogram of decision scores provides a visual representation of the model's discriminative ability. The negative class distribution is highly skewed, with a strong peak near a decision score of 0.0, indicating that the model confidently assigns low scores to the majority of samples from this class. Conversely, the positive class distribution is broader, with a noticeable concentration of scores in the high range (0.8–0.9), reflecting the model's ability to identify a large subset of positive samples with high confidence. However, the histogram also reveals a region of overlap between the two class distributions. This overlap indicates that for a portion of the samples, the model's predictive certainty is diminished, as both negative and positive class data points occupy this same range. This observation underscores the importance of considering the decision score itself, beyond a simple binary classification, especially for clinical decision support where borderline cases may require further scrutiny. Visualizations using PCA and t-SNE further illustrate the separability between classes. Although some overlap exists—as is typical in biological data—the overall clustering supports the discriminative power of HRV features. These projections also help to interpret the feature space learned by the model and highlight regions of higher classification uncertainty.

The SHAP analysis revealed that frequency domain features, particularly HF and LF, play a central role in distinguishing between IHD and non-IHD cases. These findings are consistent with physiological evidence linking reduced autonomic function, especially diminished parasympathetic activity, with cardiovascular risk [21,34,36]. The prominence of HF suggests a strong connection between vagal tone and IHD prediction.

Several limitations should be acknowledged in this pilot study. One of the key objectives was to evaluate the performance and reliability of the Zhurek IoT device, which utilizes PPG technology to collect data in real-world, ambulatory conditions. For data verification, we conducted parallel data collection using standard three-lead Holter ECG monitors. While the Holter ECG is a device that records the heart's electrical activity, allowing for the detection of rhythm disturbances, ischemia, and other issues, PPG, in turn, measures volumetric blood pulsation in peripheral vessels.

This methodological difference determines their respective strengths and weaknesses. Comparative studies confirm that ECG often outperforms other methods in terms of signal quality and respiratory signal extraction [58]. However, ECG is susceptible to interference and motion artifacts arising from unstable skin contact [59]. PPG, as an optical technique, measures volumetric changes in blood flow within the microvascular bed. Since PPG captures hemodynamic changes indirectly, factors such as poor peripheral perfusion or incorrect sensor placement (for instance, on the wrist, as with the Zhurek device) can

significantly degrade signal quality and reduce diagnostic specificity [26]. Despite this, modern research indicates that with the use of advanced signal processing algorithms, PPG-based devices can achieve performance comparable to ECG in tasks such as AF detection [60]. Our work supports this thesis, demonstrating high accuracy (AUC = 0.95) in classifying IHD based on HRV metrics obtained from PPG.

Another significant limitation of this pilot study is the small number of participants, which is particularly noticeable in the healthy control group. A small sample size can reduce the statistical power and the model's generalization ability, potentially distorting the interpretation of intergroup differences. To address the class imbalance problem, we employed a generative adversarial network, CTGAN, to create synthetic data. While synthetic data cannot fully replicate the complexity of physiological signals, it successfully approximates their statistical structure, which allowed for more balanced model training. Nevertheless, to definitively confirm the obtained results and enhance their reliability, we plan to conduct a larger-scale study with a significantly increased number of participants in both groups.

Furthermore, a critical limitation of this study is that all data were collected from participants exclusively in a resting state. While this controlled approach allowed us to establish a baseline performance and validate the device against the gold-standard Holter under ideal conditions, it does not account for the impact of motion artifacts. PPG signals are notoriously susceptible to corruption from physical activities such as walking or even minor hand movements. The current study did not evaluate the device's stability or anti-interference capabilities in dynamic, real-world scenarios. Therefore, while our results are promising for resting-state applications, the device's utility for continuous monitoring during daily life remains to be validated.

Finally, the generalizability of our findings may be limited as the model was trained and validated exclusively on a single-center sample from Kazakhstan. This specific demographic and geographic context means that the model's performance on other ethnic or regional populations is unknown. It is not possible to rule out potential biases stemming from unique genetic, lifestyle, or environmental factors characteristic of the study population. Therefore, external validation on diverse, multi-center datasets is a necessary next step to confirm the robustness and broader applicability of our model.

6. Conclusions

The pilot study showed that PPG recordings obtained with the “Zhurek” device reproduce HRV metrics with clinically acceptable accuracy. In comparative tests, the deviation from the three-lead Holter monitor was -0.601 bpm for mean HR, $+33.1$ ms for SDNN, and -4.8 ms for RMSSD. HRV data were obtained from the Research Institute of Cardiology and Internal Medicine; they were recorded using ECG Holter and the custom Zhurek IoT device. The ML models trained on HRV features showed promising results in classifying individuals with and without IHD. Among them, the RF model achieved the highest accuracy of 90.82%, demonstrating strong performance across precision (92.11%), recall (91.00%), and F1-score (90.11%) metrics. Testing conducted exclusively on real data confirmed the model's high efficacy: at the optimal cutoff threshold, the sensitivity was 88%, and the specificity was 100%. The use of CTGAN for synthetic data generation addressed the class imbalance issue by expanding the healthy control group from 50 to 250 samples, contributing to improved model generalizability and robustness. Analysis of the model's internal logic using the SHAP method showed that the contribution of each HRV feature to the prediction of IHD is modulated by changes in parasympathetic activity, as reflected in the high-frequency (HF) component. It was established that in cases of sympathovagal imbalance and low vagal tone (low HF values), metrics such as LF/HF, LF, and RMSSD

were the primary contributors to an increased risk of IHD. Conversely, in the presence of strong parasympathetic modulation (high HF values), features like SDNN, PNN50, and RMSSD played a protective role in the prediction. These results underscore that the model identified non-linear, multivariate dependencies, and that the predictive value of HRV markers is context-dependent and dynamically influenced by autonomic balance.

These findings confirm that the IoT device Zhurek is a viable and affordable platform for ambulatory HRV monitoring, and that the intelligent models developed using its data can be effectively applied to the early risk detection of IHD.

Our future work will focus on evolving the Zhurek IoT device from a validated prototype into a scalable, intelligent platform for proactive cardiovascular health management. Addressing the limitations of the current pilot study is our immediate priority.

First, and most critically, we will move our analysis from a controlled resting state to dynamic, real-world conditions. Future research will be aimed at conducting experiments to assess the device's noise immunity during various activities, such as walking and jogging. Concurrently, we will focus on developing and implementing advanced algorithms to suppress the motion artifacts inherent in PPG signals, which is essential for reliable ambulatory monitoring.

To further enhance the robustness and generalizability of our models, we will markedly increase the sample size, particularly in the healthy volunteer cohort, to improve statistical power. Critically, we plan to conduct external validation using multi-center and multi-ethnic datasets to ensure our model performs well across diverse populations and is free from the geographic or genetic biases identified in this study. We will also broaden the feature set by incorporating non-linear, geometric, and segment-based HRV indices, along with relevant clinical covariates, to create a more comprehensive picture of cardiovascular health.

The ultimate goal is to integrate the validated device with interpretable ML models. This will enable continuous ambulatory heart monitoring, provide early alerts for at-risk groups, and facilitate personalized treatment strategies. Collectively, these steps will help lessen the burden of CVDs through early intervention and proactive health management.

Author Contributions: Conceptualization, N.T., B.A., T.I. and B.I.; methodology, N.T., B.A., O.A.P. and A.K.; software, N.T., B.A. and B.I.; validation, T.I., B.I., O.A.P. and A.K.; formal analysis, T.I., B.I. and O.A.P.; investigation, N.T. and B.A.; resources, N.T., B.A., O.A.P. and A.K.; data curation, N.T., B.A., O.A.P. and A.K.; writing—original draft preparation, N.T. and B.A.; writing—review and editing, T.I., B.I. and O.A.P.; visualization, N.T. and B.A.; supervision, T.I., B.I. and O.A.P.; project administration, N.T. and B.A.; funding acquisition, T.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP26103523).

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Al-Farabi Kazakh National University (KazNU).

Informed Consent Statement: Informed consent for participation is not required as per local legislation [Protocol № IRB-A862].

Data Availability Statement: The data supporting the findings of this study are available from the corresponding authors upon request. The data are not publicly available due to privacy or ethical restrictions.

Conflicts of Interest: Author Akzhan Konysbekova was employed by the JSC “Research Institute of Cardiology and Internal Diseases”. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AF	Atrial Fibrillation
ANS	Autonomic Nervous System
AppEn	Approximate Entropy
AUC	Area Under the Curve
BMI	Body Mass Index
CAD	Coronary Artery Disease
CatBoost	Categorical Boosting Algorithm
CHF	Chronic Heart Failure
CV	Cross-Validation
CVD	Cardiovascular Disease
DNN	Deep Neural Network
DNN-LMSVM	Deep Neural Network with Least Mean Square Support Vector Machine
DT	Decision Tree
ECG	Electrocardiography
ECG Holter	Continuous Ambulatory ECG Monitoring
HRV	Heart Rate Variability
EBM	Explainable Boosting Machine
HF	High Frequency (HRV component)
HR	Heart Rate
I ² C	Inter-Integrated Circuit
IHD	Ischemic Heart Disease
JSON	JavaScript Object Notation
KNN	K-Nearest Neighbors
LF	Low Frequency (HRV component)
LF/HF	Ratio of Low to High Frequency HRV Components
LMSVM	Least Mean Square Support Vector Machine
Max_HR	Maximum Heart Rate
ML	Machine Learning
MQTT	Message Queuing Telemetry Transport
NTP	Network Time Protocol
PCA	Principal Component Analysis
PNN50	Percentage of Successive RR Intervals Differing by >50 ms
PPG	Photoplethysmography
RF	Random Forest
RMSSD	Root Mean Square of Successive Differences
ROC	Receiver Operating Characteristic
RTC	Real-Time Clock
SampEn	Sample Entropy
SD1/SD2	Poincaré Plot Standard Deviations
SDANN	Standard Deviation of the Average NN Intervals
SDNN	Standard Deviation of NN Intervals
TLS	Transport Layer Security

References

- World Health Organization. Cardiovascular Diseases (CVDs). Available online: <https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-cvds> (accessed on 11 June 2024).
- Khan, M.A.; Hashim, M.J.; Mustafa, H.; Baniyas, M.Y.; Al Suwaidi, S.; AlKatheeri, R.; Alblooshi, F.M.K.; Almatrooshi, M.; Alzaabi, M.E.H.; Al Darmaki, R.S.; et al. Global Epidemiology of Ischemic Heart Disease: Results from the Global Burden of Disease Study. *Cureus* **2020**, *12*, e9349. [CrossRef]
- Tengrinews.kz. The Most Common Disease Among Kazakhstanis Has Been Named. Available online: https://tengrinews.kz/kazakhstan_news/nazvana-samaya-rasprostranennaya-bolezn-sredi-kazahstantsev-503527/ (accessed on 18 October 2022).
- Severino, P.; D'Amato, A.; Pucci, M.; Infusino, F.; Adamo, F.; Birtolo, L.I.; Netti, L.; Montefusco, G.; Chimenti, C.; Lavalle, C.; et al. Ischemic Heart Disease Pathophysiology Paradigms Overview: From Plaque Activation to Microvascular Dysfunction. *Int. J. Mol. Sci.* **2020**, *21*, 8118. [CrossRef]
- Janicki, Ł.J.; Leoński, W.; Janicki, J.S.; Nowotarski, M.; Dziuk, M.; Piotrowicz, R. Comparative Analysis of the Diagnostic Effectiveness of SATRO ECG in the Diagnosis of Ischemia Diagnosed in Myocardial Perfusion Scintigraphy Performed Using the SPECT Method. *Diagnostics* **2022**, *12*, 297. [CrossRef]
- Duca, Ș.-T.; Tudorancea, I.; Haba, M.Ș.C.; Costache, A.-D.; Șerban, I.-L.; Pavăl, D.R.; Loghin, C.; Costache-Enache, I.-I. Enhancing Comprehensive Assessments in Chronic Heart Failure Caused by Ischemic Heart Disease: The Diagnostic Utility of Holter ECG Parameters. *Medicina* **2024**, *60*, 1315. [CrossRef] [PubMed]
- Banerjee, R.; Ghose, A.; Muthana Mandana, K. A Hybrid CNN-LSTM Architecture for Detection of Coronary Artery Disease from ECG. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 19–24 July 2020; pp. 1–8. [CrossRef]
- Ribeiro, P.; Sá, J.; Paiva, D.; Rodrigues, P.M. Cardiovascular Diseases Diagnosis Using an ECG Multi-Band Non-Linear Machine Learning Framework Analysis. *Bioengineering* **2024**, *11*, 58. [CrossRef] [PubMed]
- Gaine, S.P.; Sharma, G.; Tower-Rader, A.; Botros, M.; Kovell, L.; Parakh, A.; Wood, M.J.; Harrington, C.M. Multimodality Imaging in the Detection of Ischemic Heart Disease in Women. *J. Cardiovasc. Dev. Dis.* **2022**, *9*, 350. [CrossRef] [PubMed]
- Wang, L.; Bi, T.; Hao, J.; Zhou, T.H. Heart Diseases Recognition Model Based on HRV Feature Extraction over 12-Lead ECG Signals. *Sensors* **2024**, *24*, 5296. [CrossRef]
- Doolub, G.; Mamalakis, M.; Alabed, S.; Van der Geest, R.J.; Swift, A.J.; Rodrigues, J.C.L.; Garg, P.; Joshi, N.V.; Dastidar, A. Artificial Intelligence as a Diagnostic Tool in Non-Invasive Imaging in the Assessment of Coronary Artery Disease. *Med. Sci.* **2023**, *11*, 20. [CrossRef]
- Verma, L.; Srivastava, S. A Data Mining Model for Coronary Artery Disease Detection Using Noninvasive Clinical Parameters. *Indian J. Sci. Technol.* **2016**, *9*, 1–6. [CrossRef]
- Sayadi, M.; Varadarajan, V.; Sadoughi, F.; Chopannejad, S.; Langarizadeh, M. A Machine Learning Model for Detection of Coronary Artery Disease Using Noninvasive Clinical Parameters. *Life* **2022**, *12*, 1933. [CrossRef]
- Buccelletti, F.; Gilardi, E.; Scaini, E.; Galiuto, L.; Persiani, R.; Biondi, A.; Basile, F.; Silveri, N. Heart Rate Variability and Myocardial Infarction: Systematic Literature Review and Metanalysis. *Eur. Rev. Med. Pharmacol. Sci.* **2009**, *13*, 299–307.
- Voss, A.; Schroeder, R.; Vallverdú, M.; Schulz, S.; Cygankiewicz, I.; Vázquez, R.; Bayés de Luna, A.; Caminal, P. Short-Term vs. Long-Term Heart Rate Variability in Ischemic Cardiomyopathy Risk Stratification. *Front. Physiol.* **2013**, *4*, 364. [CrossRef]
- Agrawal, R.K.; Sewani, R.R.; Delen, D.; Benjamin, B. A Machine Learning Approach for Classifying Healthy and Infarcted Patients Using Heart Rate Variabilities Derived Vector Magnitude. *Healthc. Anal.* **2022**, *2*, 100121. [CrossRef]
- Georgieva-Tsaneva, G.; Gospodinova, E. Heart Rate Variability Analysis of Healthy Individuals and Patients with Ischemia and Arrhythmia. *Diagnostics* **2023**, *13*, 2549. [CrossRef]
- Georgieva-Tsaneva, G.; Cheshmedzhiev, K.; Tsanev, Y.-A.; Dechev, M.; Popovska, E. Healthcare Monitoring Using an Internet of Things-Based Cardio System. *IoT* **2025**, *6*, 10. [CrossRef]
- Alaa, A.M.; Bolton, T.; Di Angelantonio, E.; Rudd, J.H.F.; van der Schaar, M. Cardiovascular Disease Risk Prediction Using Automated Machine Learning: A Prospective Study of 423,604 UK Biobank Participants. *PLoS ONE* **2019**, *14*, e0213653. [CrossRef]
- Liu, W.; Xin, Y.; Sun, M.; Liu, C.; Yin, X.; Xu, X.; Xiao, Y. Relationship Between Heart Rate Variability Traits and Stroke: A Mendelian Randomization Study. *J. Stroke Cerebrovasc. Dis.* **2025**, *34*, 108251. [CrossRef] [PubMed]
- Aftyka, J.; Staszewski, J.; Dębiec, A.; Pogoda-Wesołowska, A.; Kowalska, A.; Jankowska, A.; Żebrowski, J. The Hemisphere of the Brain in Which a Stroke Has Occurred Visible in the Heart Rate Variability. *Life* **2022**, *12*, 1659. [CrossRef] [PubMed]
- Alizadehsani, R.; Hosseini, M.J.; Khosravi, A.; Khozeimeh, F.; Roshanzamir, M.; Sarrafzadegan, N.; Nahavandi, S. Non-Invasive Detection of Coronary Artery Disease in High-Risk Patients Based on the Stenosis Prediction of Separate Coronary Arteries. *Comput. Methods Programs Biomed.* **2018**, *162*, 119–127. [CrossRef] [PubMed]
- Brinza, C.; Floria, M.; Scripcariu, D.-V.; Covic, A.M.; Covic, A.; Popa, I.V.; Statescu, C.; Burlacu, A. Heart Rate Variability in Acute Myocardial Infarction: Results of the HeaRt-V-AMI Single-Center Cohort Study. *J. Cardiovasc. Dev. Dis.* **2024**, *11*, 254. [CrossRef]

24. Hazra, A.; Mandal, S.K.; Gupta, A.; Mukherjee, A.; Mukherjee, A. Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review. *Adv. Comput. Sci. Technol.* **2017**, *10*, 2137–2159.
25. Moraes, J.L.; Rocha, M.X.; Vasconcelos, G.G.; Vasconcelos Filho, J.E.; De Albuquerque, V.H.C.; Alexandria, A.R. Advances in Photoplethysmography Signal Analysis for Biomedical Applications. *Sensors* **2018**, *18*, 1894. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Elgendi, M.; Fletcher, R.; Liang, Y.; Howard, N.; Lovell, N.H.; Abbott, D.; Lim, K.; Ward, R. The Use of Photoplethysmography for Assessing Hypertension. *npj Digit. Med.* **2019**, *2*, 60. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Almarshad, M.A.; Islam, M.S.; Al-Ahmadi, S.; BaHammam, A.S. Diagnostic Features and Potential Applications of PPG Signal in Healthcare: A Systematic Review. *Healthcare* **2022**, *10*, 547. [\[CrossRef\]](#)
28. Kim, K.B.; Baek, H.J. Photoplethysmography in Wearable Devices: A Comprehensive Review of Technological Advances, Current Challenges, and Future Directions. *Electronics* **2023**, *12*, 2923. [\[CrossRef\]](#)
29. Shabaan, M.; Arshid, K.; Yaqub, M.; Masud, M.; AlZain, M.A.; Shahzad, H.; Rodrigues, J.J.P.C. Survey: Smartphone-Based Assessment of Cardiovascular Diseases Using ECG and PPG Analysis. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 177. [\[CrossRef\]](#)
30. Jahmunah, V.; Ng, E.Y.K.; San, T.R.; Acharya, U.R. Automated Detection of Coronary Artery Disease, Myocardial Infarction and Congestive Heart Failure Using GaborCNN Model with ECG Signals. *Comput. Biol. Med.* **2021**, *134*, 104457. [\[CrossRef\]](#)
31. Eom, G.; Byeon, H. Searching for Optimal Oversampling to Process Imbalanced Data: Generative Adversarial Networks and Synthetic Minority Over-Sampling Technique. *Mathematics* **2023**, *11*, 3605. [\[CrossRef\]](#)
32. Ogunpola, A.; Saeed, F.; Basurra, S.; Albarrak, A.M.; Qasem, S.N. Machine Learning-Based Predictive Models for Detection of Cardiovascular Diseases. *Diagnostics* **2024**, *14*, 144. [\[CrossRef\]](#)
33. Aftyka, J.; Staszewski, J.; Debiec, A.; Pogoda-Wesołowska, A.; Żebrowski, J. Can HRV Predict Prolonged Hospitalization and Favorable or Unfavorable Short-Term Outcome in Patients with Acute Ischemic Stroke? *Life* **2023**, *13*, 856. [\[CrossRef\]](#)
34. Chairina, G.; Yoshino, K.; Kiyono, K.; Watanabe, E. Ischemic Stroke Risk Assessment by Multiscale Entropy Analysis of Heart Rate Variability in Patients with Persistent Atrial Fibrillation. *Entropy* **2021**, *23*, 918. [\[CrossRef\]](#)
35. Marcantoni, I.; Iammarino, E.; Dell’Orletta, A.; Burattini, L. Prognostic Role of Electrocardiographic Alternans in Ischemic Heart Disease. *J. Clin. Med.* **2025**, *14*, 2620. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Kaufmann, D.K.; Raczak, G.; Szwoch, M.; Wabich, E.; Świątczak, M.; Daniłowicz-Szymanowicz, L. Baroreflex Sensitivity but Not Microvolt T-Wave Alternans Can Predict Major Adverse Cardiac Events in Ischemic Heart Failure. *Cardiol. J.* **2022**, *29*, 1004–1012. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Kusunose, K.; Abe, T.; Haga, A.; Fukuda, D.; Yamada, H.; Harada, M.; Sata, M. A Deep Learning Approach for Assessment of Regional Wall Motion Abnormality from Echocardiographic Images. *JACC Cardiovasc. Imaging* **2020**, *13*, 374–381. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Zreik, M.; Lessmann, N.; van Hamersvelt, R.W.; Wolterink, J.M.; Voskuil, M.; Viergever, M.A.; Išgum, I. Deep Learning Analysis of Coronary Arteries in Cardiac CT Angiography for Detection of Patients Requiring Invasive Coronary Angiography. *IEEE Trans. Med. Imaging* **2020**, *39*, 1545–1557. [\[CrossRef\]](#)
39. Ripan, R.C.; Sarker, I.H.; Hossain, S.M.M.; Ahmmed, R.; Ghosh, S.; Uddin, M. A Data-Driven Heart Disease Prediction Model Through K-Means Clustering-Based Anomaly Detection. *SN Comput. Sci.* **2021**, *2*, 112. [\[CrossRef\]](#)
40. Dipto, I.; Islam, T.; Rahman, H.; Rahman, M. Comparison of Different Machine Learning Algorithms for the Prediction of Coronary Artery Disease. *J. Data Anal. Inf. Process.* **2020**, *8*, 41–68. [\[CrossRef\]](#)
41. Trigka, M.; Dritsas, E. Long-Term Coronary Artery Disease Risk Prediction with Machine Learning Models. *Sensors* **2023**, *23*, 1193. [\[CrossRef\]](#)
42. Wu, M.-J.; Dewi, S.R.K.; Hsu, W.-T.; Hsu, T.-Y.; Liao, S.-F.; Chan, L.; Lin, M.-C. Exploring Relationships of Heart Rate Variability, Neurological Function, and Clinical Factors with Mortality and Behavioral Functional Outcome in Patients with Ischemic Stroke. *Diagnostics* **2024**, *14*, 1304. [\[CrossRef\]](#)
43. Birn, R.M.; Cornejo, M.D.; Molloy, E.K.; Patriat, R.; Meier, T.B.; Kirk, G.R.; Nair, V.A.; Meyerand, M.E.; Prabhakaran, V. The Influence of Physiological Noise Correction on Test–Retest Reliability of Resting-State Functional Connectivity. *Brain Connect.* **2014**, *4*, 511–522. [\[CrossRef\]](#)
44. Sanchez-Delgado, Y.; Alcantara, J.M.A.; Ortiz-Alvarez, L.; Xu, H.; Martinez-Tellez, B.; Labayen, I.; Ruiz, J.R. Reliability of Resting Metabolic Rate Measurements in Young Adults: Impact of Methods for Data Analysis. *Clin. Nutr.* **2018**, *37*, 1618–1624. [\[CrossRef\]](#) [\[PubMed\]](#)
45. McDuff, D.; Gontarek, S.; Picard, R. Remote Measurement of Cognitive Stress via Heart Rate Variability. In Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, USA, 26–30 August 2014; pp. 2957–2960. [\[CrossRef\]](#)
46. Taoum, A.; Bisiaux, A.; Tilquin, F.; Le Guillou, Y.; Carrault, G. Validity of Ultra-Short-Term HRV Analysis Using PPG—A Preliminary Study. *Sensors* **2022**, *22*, 7995. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Hoog Antink, C.; Mai, Y.; Peltokangas, M.; Leonhardt, S.; Oksala, N.; Vehkaoja, A. Accuracy of Heart Rate Variability Estimated with Reflective Wrist-PPG in Elderly Vascular Patients. *Sci. Rep.* **2021**, *11*, 8123. [\[CrossRef\]](#) [\[PubMed\]](#)

48. Roerecke, M.; Rehm, J. Alcohol Consumption, Drinking Patterns, and Ischemic Heart Disease: A Narrative Review of Meta-Analyses and a Systematic Review and Meta-Analysis of the Impact of Heavy Drinking Occasions on Risk for Moderate Drinkers. *BMC Med.* **2014**, *12*, 182. [\[CrossRef\]](#)
49. Ng, R.; Sutradhar, R.; Yao, Z.; Wodchis, W.P.; Rosella, L.C. Smoking, Drinking, Diet and Physical Activity—Modifiable Lifestyle Risk Factors and Their Associations with Age to First Chronic Disease. *Int. J. Epidemiol.* **2020**, *49*, 113–130. [\[CrossRef\]](#)
50. Perulli, M.; Scala, I.; Venditti, R.; Amadio, A.; Gambardella, M.L.; Quintiliani, M.; Contaldo, I.; Veredice, C.; Della Marca, G.; Brunetti, V.; et al. Short- vs Long-Term Assessment of Heart Rate Variability: Clinical Significance in Dravet Syndrome. *Epilepsy Behav.* **2023**, *146*, 109357. [\[CrossRef\]](#)
51. Vandenberk, B.; Floré, V.; Röver, C.; Vos, M.A.; Dunnink, A.; Leftheriotis, D.; Friede, T.; Flevari, P.; Zabel, M.; Willems, R. Repeating Noninvasive Risk Stratification Improves Prediction of Outcome in ICD Patients. *Ann. Noninvasive Electrocardiol.* **2020**, *25*, e12794. [\[CrossRef\]](#)
52. Oehr, P. Interrelationships Among Sensitivity, Precision, Accuracy, Specificity and Predictive Values in Bioassays, Represented as Combined ROC Curves with Integrated Cutoff Distribution Curves and Novel Index Values. *Diagnostics* **2025**, *15*, 410. [\[CrossRef\]](#)
53. Steinberg, C.; Philippon, F.; Sanchez, M.; Fortier-Poisson, P.; O'Hara, G.; Molin, F.; Sarrazin, J.-F.; Nault, I.; Blier, L.; Roy, K.; et al. A Novel Wearable Device for Continuous Ambulatory ECG Recording: Proof of Concept and Assessment of Signal Quality. *Biosensors* **2019**, *9*, 17. [\[CrossRef\]](#)
54. Freund, O.; Caspi, I.; Shacham, Y.; Frydman, S.; Biran, R.; Abu Katash, H.; Zornitzki, L.; Bornstein, G. Holter ECG for Syncope Evaluation in the Internal Medicine Department—Choosing the Right Patients. *J. Clin. Med.* **2022**, *11*, 4781. [\[CrossRef\]](#)
55. Kim, H.; Huh, K.Y.; Piao, M.; Ryu, H.; Yang, W.; Lee, S.; Kim, K.H. Self-Reporting Technique-Based Clinical-Trial Service Platform for Real-Time Arrhythmia Detection. *Appl. Sci.* **2022**, *12*, 4558. [\[CrossRef\]](#)
56. Guede-Fernández, F.; Ferrer-Mileo, V.; Mateu-Mateus, M.; Ramos-Castro, J.; García-González, M.Á.; Fernández-Chimeno, M. A Photoplethysmography Smartphone-Based Method for Heart Rate Variability Assessment: Device Model and Breathing Influences. *Biomed. Signal Process. Control* **2020**, *57*, 101717. [\[CrossRef\]](#)
57. Adiputra, I.N.M.; Wanchai, P. CTGAN-ENN: A Tabular GAN-Based Hybrid Sampling Method for Imbalanced and Overlapped Data in Customer Churn Prediction. *J. Big Data* **2024**, *11*, 121. [\[CrossRef\]](#)
58. Orphanidou, C. Derivation of Respiration Rate from Ambulatory ECG and PPG Using Ensemble Empirical Mode Decomposition: Comparison and Fusion. *Comput. Biol. Med.* **2016**, *81*, 45–54. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Charlton, P.H.; Birrenkott, D.A.; Bonnici, T.; Pimentel, M.A.F.; Johnson, A.E.W.; Alastruey, J.; Beale, R.; Watkinson, P.J.; Tarassenko, L.; Clifton, D.A. Breathing Rate Estimation from the Electrocardiogram and Photoplethysmogram: A Review. *IEEE Rev. Biomed. Eng.* **2018**, *11*, 2–20. [\[CrossRef\]](#)
60. Gruwez, H.; Evens, S.; Proesmans, T.; Duncker, D.; Linz, D.; Heidbuchel, H.; Manninger, M.; Vandervoort, P.; Haemers, P.; Pison, L. Accuracy of Physicians Interpreting Photoplethysmography and Electrocardiography Tracings to Detect Atrial Fibrillation: INTERPRET-AF. *Front. Cardiovasc. Med.* **2021**, *8*, 734737. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.