

Revisiting Left Ventricular Ejection Fraction Levels: A Circadian Heart Rate Variability-based Approach

Mohanad Alkhodari (✉ mohanad.alkhodari@ku.ac.ae)

Khalifa University

Herbert F. Jelinek

Khalifa University

Shiza Saleem

Khalifa University

Leontios J. Hadjileontiadis

Khalifa University

Ahsan H. Khandoker

Khalifa University

Research Article

Keywords: heart failure, clinical practice, coronary artery disease (CAD), left ventricular ejection fraction (LVEF), heart rate variability (HRV)

Posted Date: March 31st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-367988/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Revisiting left ventricular ejection fraction levels: a circadian heart rate variability-based approach

Mohanad Alkhodari^{1,*}, Herbert F. Jelinek^{1,2,3}, Shiza Saleem¹, Leontios J. Hadjileontiadis^{1,3,4}, and Ahsan H. Khandoker^{1,3}

¹Healthcare Engineering Innovation Center (HEIC), Department of Biomedical Engineering, Khalifa University, Abu Dhabi, UAE.

²Biotechnology Center (BTC), Department of Biomedical Engineering, Khalifa University, Abu Dhabi, UAE.

³Department of Electrical Engineering and Computer Science, Khalifa University, Abu Dhabi, UAE.

⁴Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece.

*Corresponding author. Email: mohanad.alkhodari@ku.ac.ae

ABSTRACT

Analysis of heart failure is important in clinical practice to ensure coronary artery disease (CAD) patients will be provided with appropriate timely treatment. The current gold standard, echocardiography, although reliable, provides a once-off left ventricular ejection fraction (LVEF) measurement but does not provide information about heart function during day / night cardiac cycles. The aim of this study was to investigate the ability of heart rate variability (HRV) features in classifying CAD patients into multiple LVEF groups through a new index, HRV ejection fraction (HRVEF) and its agreement to established norms. A total of 92 heart failure patients were included in this study. HRVEF groups optimized on hourly basis through Jenks natural breaks algorithm exhibited a consistent pattern with a goodness of variance fit (GVF) of more than 0.70 accuracy during the late-night to early-morning (01:00-08:00) and evening (17:00-23:00) time periods. At these hours, several HRV features were found significant in differentiating between HRVEF groups using statistical analysis of variance (ANOVA) test. These features include the successive differences between normal heartbeats (RMSSD), low and high frequency (LF, HF) power, standard deviation of normal heartbeats (SD2), short-term scaling exponent (α_1), and percentage of normal heartbeats in alternation segments (PAS). This study paves the way towards new LVEF guidelines based on the association between LVEF and HRV features for a better demonstration of the circadian cardiac function at different LVEF levels in CAD patients.

Introduction

Heart failure is a chronic condition that is characterized by damaged or a weakened heart muscle that often leads to a reduction in the overall cardiac output¹. World-wide, it is estimated that more than 26 million people are suffering from heart failure². Early diagnosis as well as timely medication helps people to live a relatively normal life³. Left ventricular ejection fraction (LVEF) is the main clinical indicator in the diagnosis of heart failure and reflects the amount of blood leaving the left ventricle at each contraction^{3,4}.

The American College of Cardiology Foundation and the American Heart Association (ACCF/AHA)⁵ classifies heart failure based on LVEF into three main groups; namely heart failure with preserved EF (HFpEF, $EF > 50\%$), heart failure with reduced EF (HFrEF, $EF < 40\%$), and heart failure with mid-range EF (HFmEF, $40\% \leq EF \leq 50\%$). In contrast, the American Society of Echocardiography and the European Association of Cardiovascular Imaging (ASE/EACVI)⁶ define HFpEF as $> 55\%$, HFrEF to $< 50\%$, and HFmEF ranging between ($50\% \leq EF \leq 55\%$). Clinically, the current gold standard for measuring LVEF is echocardiography⁷. Despite being an efficient technique in determining LVEF, it is considered expensive and not often available in public healthcare services⁸. In addition, it is a once-off measurement for cardiac output, therefore, it is not able to provide time variant measurements as well as an overall representation of the heart rhythm with respect to cardiac output throughout the day/night cycles. Additionally, although the current ACCF/AHA and ASE/EACVI guidelines provide classifications for patients based on LVEF levels measured using echocardiography, there is disagreement between cut-off values for preserved, mid-range, and reduced patient groups. Patients from these groups may exhibit similar heart functionality at the time of the echocardiography test. This usually occurs for the HFmEF patients group that despite of their lower LVEF, a normal heart function is still observed and hence this group is usually considered ambiguous. Therefore, a progression of heart failure may arise in patients classified with normal or close to normal heart function.

Developing alternative indicators for heart failure associated with left ventricular systolic dysfunction is an important clinical aim. One such option is the use of electrocardiography (ECG) and the corresponding heart rate variability (HRV) features⁹. HRV represents the changes in cardiac interbeat (RR) intervals that may be associated with endocrine, autonomic

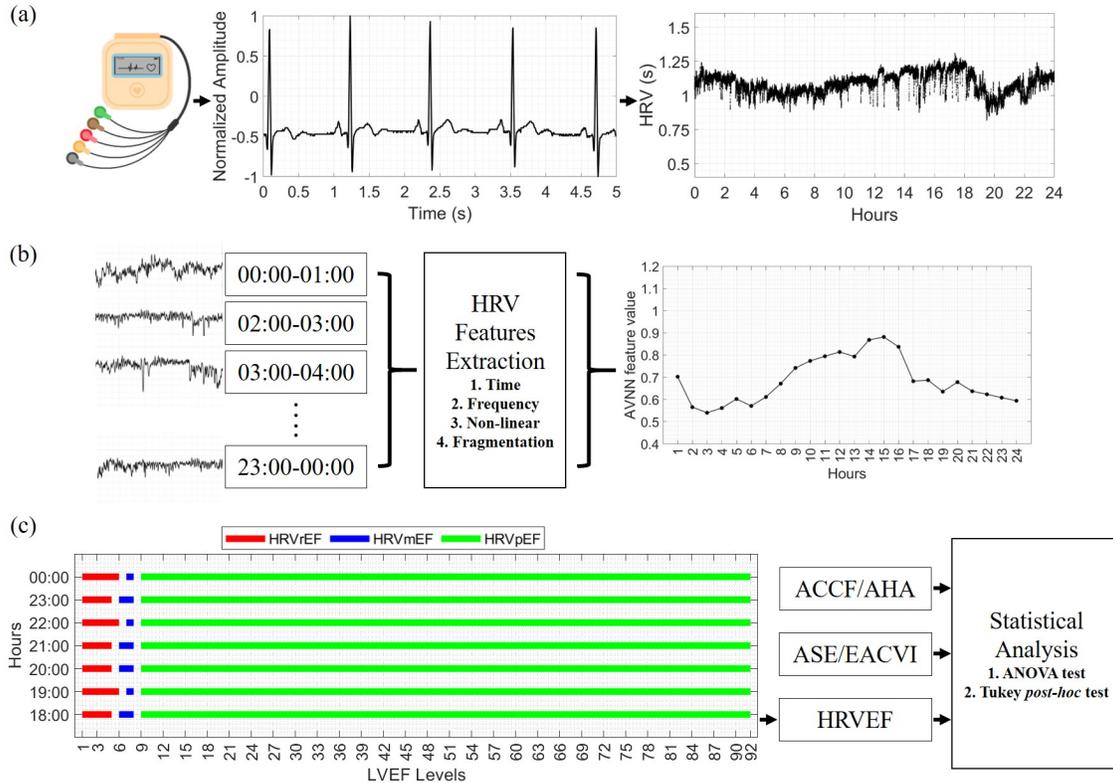


Figure 1. Flowchart for the generation of per-hour heart rate variability ejection fraction (HRVEF) index. (a) HRV acquisition from 24-hour Holter ECG recordings of coronary artery disease (CAD) patients. (b) Per-hour extraction of HRV features from time, frequency, non-linear, and fragmentation metrics (showing average normal-to-normal feature (AVNN)). (c) Utilizing Jenks natural breaks algorithm to categorize patients with respect to their left ventricular ejection fraction (LVEF) levels and statistically comparing them with gold standard guidelines.

nervous system (ANS), or intrinsic modulation of cardiac rhythm¹⁰. In coronary artery disease (CAD) patients, changes in the autonomic regulatory balance involves both the sympathetic and parasympathetic branches depending on the underlying cause and disease progression¹¹. Despite the wide usage of HRV in cardiovascular disease analysis, a deep understanding of the relationship between LVEF and HRV is still lacking and is not well defined. Further investigations on the mechanistic functionality of the heart as observed by HRV features is required in the analysis of heart failure progression in CAD patients. A promising alternative to the current gold standard (echocardiography) is HRV, which provides a more extensive perspective about the functionality of the heart at different LVEF levels¹²⁻¹⁴.

Several studies have observed a strong association between HRV and cardiovascular diseases including CAD¹⁵⁻²¹. Specific cardiac function characteristics observed using different HRV features at certain times of the day/night cycle could suggest changes in cardiac function associated with heart failure progression in parallel with circadian rhythm changes. Thus, HRV ejection fraction (HRVEF) levels can be considered as multi-time measurements as opposed to the once-off measurements provided by echocardiography. Thus, a combination of HRV, which has been shown to be a strong indicator for sudden cardiac death, combined with LVEF may provide a more powerful assessment tool for treatment and medication options for patients with established heart failure.

The current study (Figure 1) aimed to investigate the ability of HRV and its corresponding features to classify a total of 92 CAD patients into LVEF groups as defined by ACCF/AHA and ASE/EACVI. Unlike the current gold standard, the new index, HRVEF, is based on optimizing group membership into HFpEF, HFmEF, and HFrfEF using HRV derived time-domain, frequency-domain, non-linear, and fragmentation metrics. The optimization was based on Jenks natural breaks algorithm to determine which HRV features and cut-off values provide the best fit into HFpEF, HFmEF, and HFrfEF classes. The advantage of using this approach lies in segmenting LVEF data in 1-dimension (1D) based on the values of features obtained from every patient at every hour over a 24-hour interval. Therefore, in addition to the current ACCF/AHA and ASE/EACVI LVEF classification guidelines, the HRVEF grouping may provide a more comprehensive per-hour classifications that combines LVEF and HRV features.

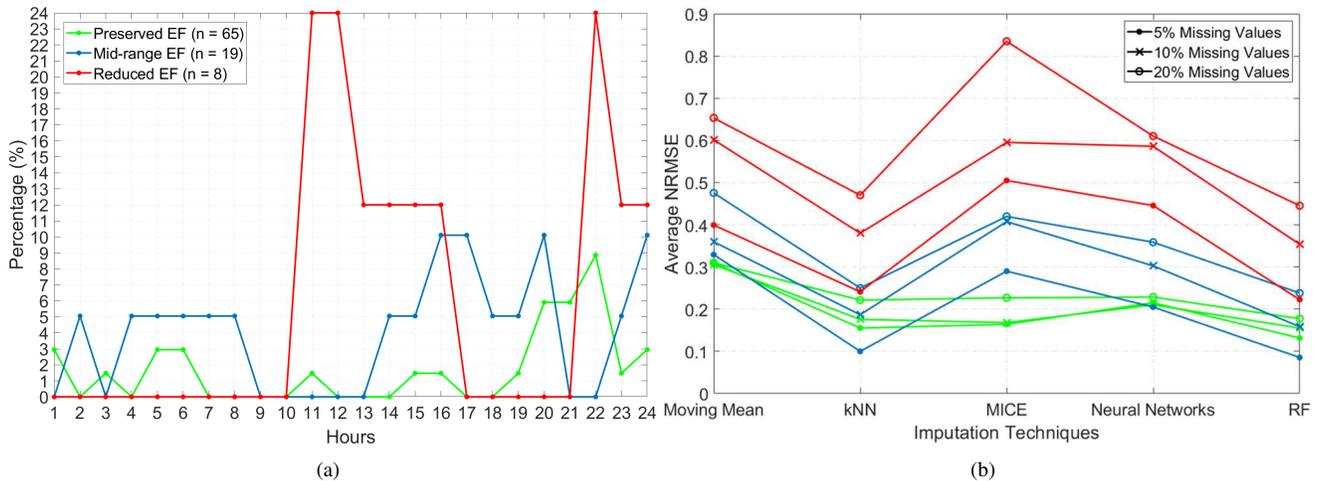


Figure 2. HRV feature imputation process to ensure a complete features dataset with no missing values. (a) Percentage of missing values across the 24-hour HRV data for each LVEF patient groups. (b) Average normalized root mean square error (NRMSE) for the five data imputation techniques. Three missing values scenarios (Dot: 5%, Cross: 10%, and Circle: 20%) were applied on the preserved (green), mid-range (blue), and reduced (red) LVEF groups to evaluate each technique.

59 Results

60 Missing HRV features imputation

61 Initially, missing HRV features, as a result of a corrupted recording or an unrecorded hour, on each hour were imputed for the
 62 three LVEF groups. An example to show missing values percentages was taken from the data set for the three LVEF groups
 63 based on ACCF/AHA guidelines (Fig. 2(a)). This example roughly shows three ranges for the missing values, namely 5%,
 64 10%, and 20% across the three groups. These missing values were considered to be missing completely at random (MCAR),
 65 which means that there are no relationships (dependency) between features across each patient. It worth noting that the reduced
 66 LVEF group had the least number of patients data (8), therefore, missing values percentage was often higher than other groups.
 67 In addition, hour 9:00 AM had a complete dataset for the three groups, thus, it was selected for further analysis.

68 To adjust the dataset to include complete 24-hour HRV features data, several data imputation techniques were applied
 69 (discussed in Section) based on conventional or sophisticated methods. Among these techniques are the moving mean filling,
 70 k-nearest neighbour (kNN), multiple imputation by chained equation (MICE), neural networks, and random forest (RF). To
 71 decide on which technique to use for HRV missing data imputation, a systematic approach was followed. Initially, an amount of
 72 5%, 10%, and 20% of missing values was introduced to per-hour patients' data from each LVEF group. Then, missing values
 73 were imputed using each technique and evaluated through the normalized root mean squared error (NRMSE) between the
 74 original and the newly imputed dataset.

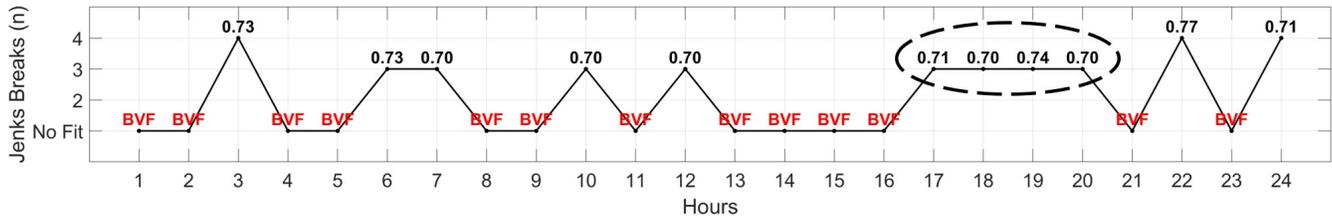
75 The performance of the data imputation techniques is depicted in Fig. 2(b) for the preserved (green), mid-range (blue),
 76 and reduced (red) LVEF groups classified based on the ACCF/AHA guidelines. The figure shows the average NRMSE value
 77 using each imputation technique with 5%, 10%, and 20% missing values scenarios. Based on the figure, kNN and RF methods
 78 had the lowest error among the three LVEF groups. For the 5% missing values scenario, kNN had error values of 0.16, 0.10,
 79 and 0.24 for the preserved, mid-range, and reduced LVEF groups, respectively. On the other hand, the error for these three
 80 groups was slightly lower using RF with 0.13, 0.08, and 0.22. For the 10% missing values scenario, the error increased to 0.18,
 81 0.19, and 0.38 using kNN and 0.15, 0.15, and 0.35 using RF for the three groups, respectively. The highest error values were
 82 observed for the 20% missing values scenario with 0.22, 0.25, and 0.47 for kNN and 0.17, 0.23, and 0.44 for RF. The highest
 83 NRMSE values were observed for the reduced LVEF group across all imputation techniques. This is due to the small number of
 84 samples (8) in the reduced group as compared to the preserved (65) and mid-range (19) groups.

85 Based on the observation obtained from the average NRMSE of the five data imputation techniques, RF was selected to
 86 impute the HRV data set. The average error values across the three missing values scenarios for preserved, mid-range, and
 87 reduced groups were 0.15, 0.15, and 0.34, respectively. These values were less than kNN which had values of 0.19, 0.18,
 88 and 0.36 for preserved, mid-range, and reduced groups, respectively. The selection was not only based on the best average NRMSE
 89 values, but also was due to the ability of RF as a machine learning method to learn from all other features prior to predicting the
 90 missing value in the data set.

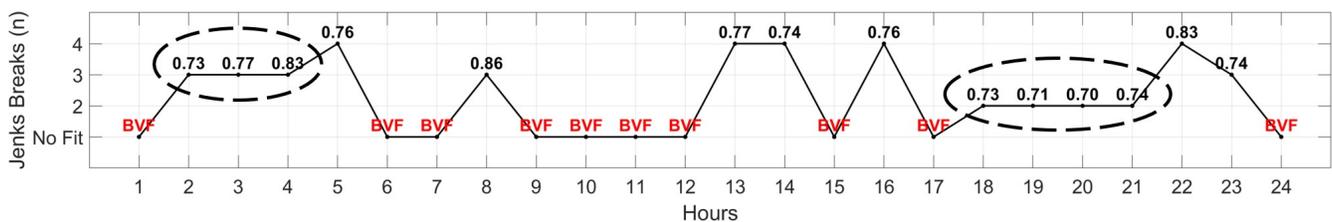
91 **Generating HRVEF groups**

92 **Natural breaks optimization**

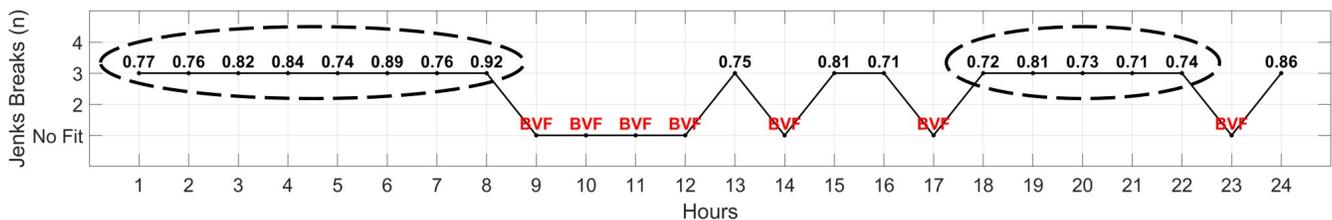
93 The outcomes of the Jenks natural break optimization algorithm in generating HRVEF groups is shown in Fig. 3 for a feature
 94 from every HRV metric including RMSSD, HF Peak, SD2, and PAS features. These features have exhibited a unique pattern
 95 in the number of Jenks breaks at specific time periods more than other features. These time periods were mostly during the



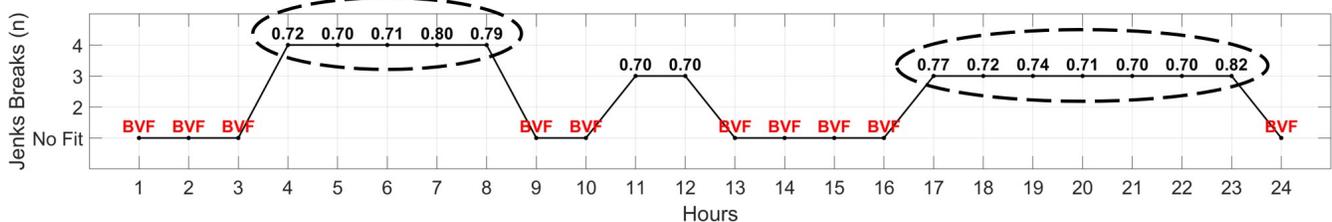
(a) Time-domain: RMSSD



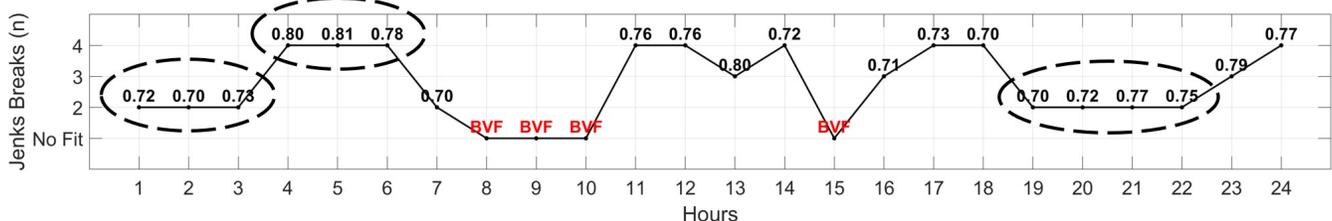
(b) Frequency-domain: HF Peak



(c) Frequency-domain: LF Power



(d) Non-linear: SD2



(e) Fragmentation: PAS

Figure 3. Heart rate variability ejection fraction (HRVEF) grouping using Jenks for selected HRV features. The figure shows the number of Jenks breaks whenever the goodness of variance fit (GVF) have reached ≥ 0.70 for (a) time-domain: RMSSD. (b) Frequency-domain: HF Peak. (c) Frequency-domain: LF Power. (d) Non-linear: SD2. (e) Fragmentation: PAS. Bad variance fit (BVF) was assigned for $GVF < 0.70$. The dashed circles denote having three or more consecutive ≥ 0.70 GVF.

96 late-night to early-morning (01:00-08:00) and evening (17:00-23:00) hours, where three or more consecutive GVF values
 97 of more than 0.70 were observed. The complete Jenks natural breaks optimization for every HRV features is provided in
 98 Supplementary Fig. S1.

99 From the time-domain, RMSSD had GVF values ranging between 0.70 to 0.74 with three natural breaks during the evening
 100 hours. The HF Peak feature from the frequency-domain had close pattern during the evening but with two natural breaks.
 101 In addition, three breaks had a better fitting performance during the late-night hours with values in the range of [0.73-0.83].

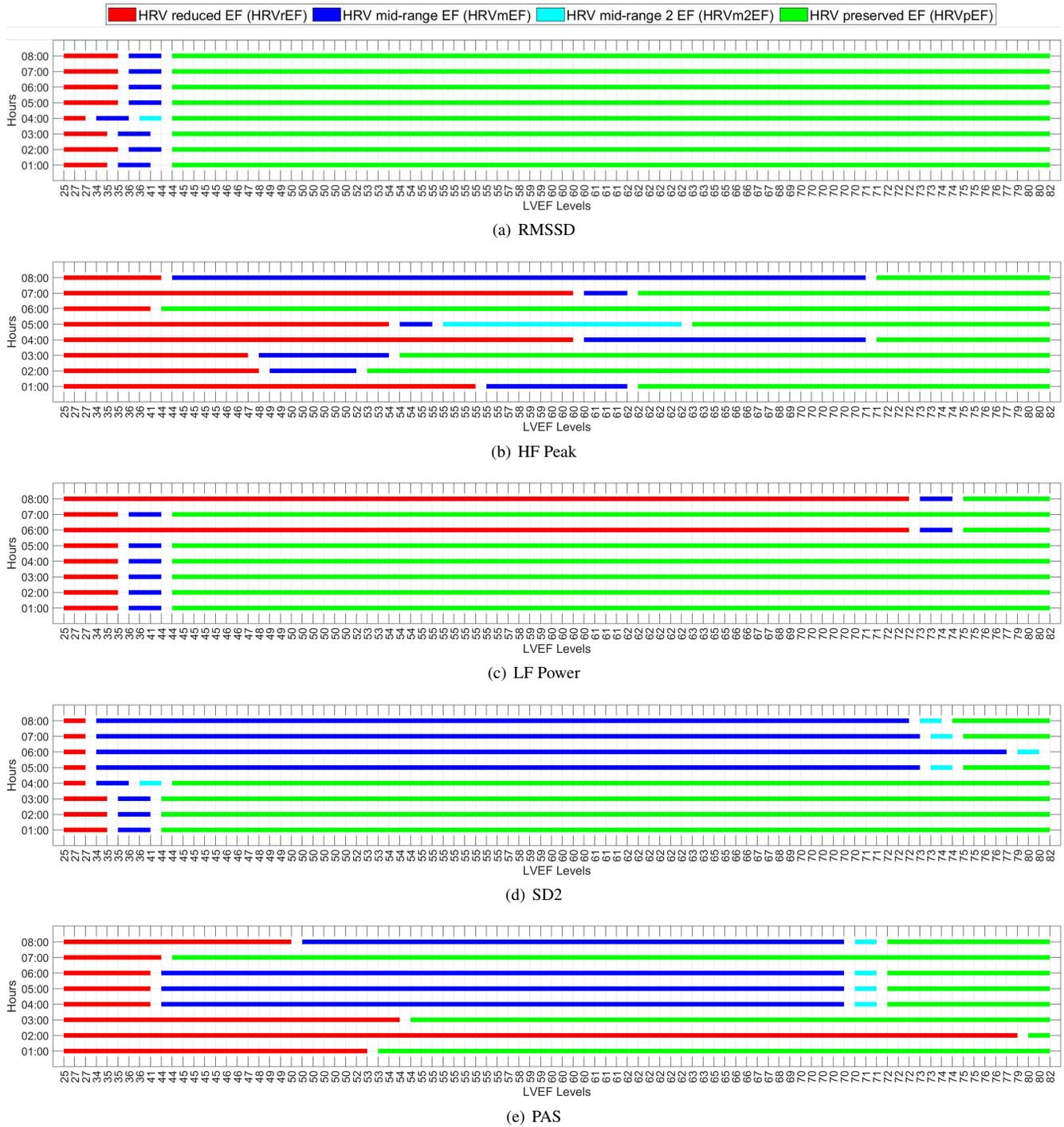


Figure 4. The late-night (01:00-03:00) to early-morning (04:00-08:00) hours distribution of heart rate variability ejection fraction (HRVEF) groups across patients based on Jenks natural breaks.

102 Furthermore, LF Power returned three natural breaks during the late-night to morning and evening hours with GVF values
 103 of more than 0.70. For the non-linear metrics, SD2 had the most continuous pattern with four natural breaks during the
 104 early-morning hours and three natural breaks during the evening hours. For the fragmentation metrics, PAS had the best
 105 performance with two breaks during the late-night and evening hours and four groups during the early-morning hours. A GVF
 106 of more than 0.80 was observed mainly when using four natural breaks.

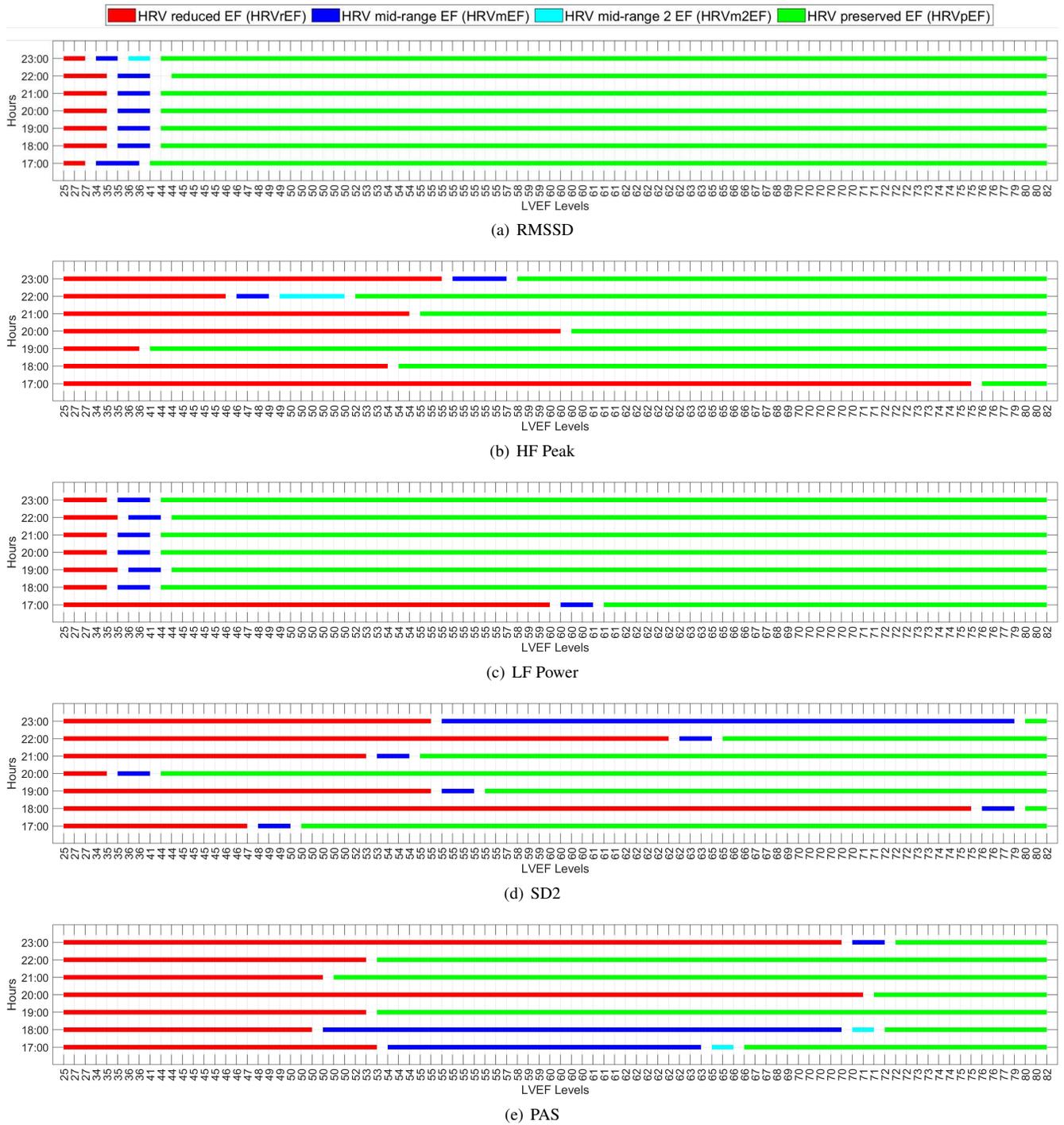


Figure 5. The evening (17:00-23:00) hours distribution of heart rate variability ejection fraction (HRVEF) groups across patients based on Jenks natural breaks.

HRVEF groups distribution

The distribution of HRVEF groups based on Jenks natural breaks is depicted in Fig. 4 and 5 for the late-night to early-morning (01:00-08:00) and evening (17:00-23:00) hours, respectively. The complete 24-hour spectrum for every HRV features is provided in Supplementary Fig. S2.

During the late-night to early-morning hours, RMSSD had a close pattern for groups distribution. The HRVpEF group was always determined to be more than 44% LVEF. Furthermore, HRVmEF group ranged between 36% and 44% LVEF percent levels except for hours 01:00, 03:00, and 04:00 where an additional HRVm2EF group was observed. The HRVrEF group ranged between 25% and 35% LVEF levels for almost all the hours. The HF Peak feature had more variability in the decision for groups limits across the hours. The late-night hours (01:00-04:00) suggested three groups with varying HRVmEF group limits of more than 48% LVEF. On the other hand, the early morning hours (05:00-06:00) had four and two groups, respectively. The HRVpEF had a wider range across most hours. Similarly to RMSSD, LF Power had three groups estimation during most of the hours with the HRVmEF group ranging between 36% to 44% except for two early-morning hours where the majority of patients were considered in the HRVrEF group. SD2 feature showed a unique pattern for the HRVmEF groups. The late-night hours had three groups scenario with HRVmEF limits between 35% and 41% LVEF levels. The early-morning hours had four groups with an additional HRVm1EF at high LVEF levels ranging between 73% and 74%. However, the majority of patients had similar HRV characteristics, thus, they were included in the HRVmEF group ranging between 34% and 73%. The last HRV feature, PAS, suggested a two group scenario during the late-night hours with a LVEF range between 25% and 54% for HRVrEF and 54% to 82% for HRVpEF. In addition, early-morning hours had four grouping scenario with the HRVmEF group ranging between 44% and 70% and HRVm1EF ranging between 70% and 71%.

During the evening hours, the distribution of HRFEV groups had close patterns to the one in the late-night to early morning hours using the RMSSD feature. The range of the HRVmEF group was between 34% to 41% for most hours except for hour 22:00-23:00 where an additional HRVm2EF group was needed for patients with 35% to 41% LVEF levels. The HRVpEF had patients with LVEF levels of more than 41% while the HRVrEF had patients with levels of less than 34%. For the HF Peak feature, two groups were essential to divide patients during most hours without exhibiting a unique pattern. Most patients were considered as HRVrEF with levels of less than 55%, while HRVpEF patients were suggested to have more than 55% except for hours 16:00-17:00 and 18:00-19:00. A pattern close to RMSSD was observed for the LF Power feature with three groups at most hours. The range for HRVmEF group was between 35% and 44% with values less than 35% for the HRVrEF group and more than 44% for the HRVpEF group. The SD2 feature had three groups scenario across all hours with varying HRVmEF limits. However, it is worth noting that hours 22:00-23:00 covered a wider range of patients within these groups with LVEF values between 55% and 79%. The last feature, PAS, suggested four groups scenario during the 16:00-18:00 time window with HRVm1EF of more than 65% and less than 71%. However, most hour intervals had two group scenarios with HRVpEF patients presenting with LVEF levels of more than 53%.

Statistically discriminative features

Table 1 shows the values of features during two time periods, early-morning (03:00-04:00) and evening (18:00-19:00), where most significant differences occurred across features. The table shows feature values (Mean \pm STD) for ACCF/AHA, ASE/EACVI, and HRV/Jenks guidelines.

Based on ACCF/AHA guidelines, RMSSD showed significant differences between the groups during the late-night (02:00-03:00) and evening (17:00-22:00). In addition, HF Power and LF Power showed similar patterns during the same time periods. However, Alpha1 was only significantly different between 05:00-06:00. On the other hand, following the ASE/EACVI guidelines, RMSSD was significant during the late-night period (02:00-04:00) and bewtwwen 21:00-22:00. Similarly to ACCF/AHA guidelines, HF Power had a similar pattern in the distribution of significant differences across the 24 hours. However, LF Power showed significant differences during the early-morning period (03:00-06:00 and 10:00-11:00 and evening (18:00-19:00). For Alpha1, a significant difference occurred across all hours due to the drop in the values in the HFrEF group.

Compared to ACCF/AHA and ASE/EACVI, using HRVEF groups, RMSSD, HF Power, LF Power, and Alpha1 have exhibited close patterns in their significant differences across hours. This pattern was mostly focused on the late-night (00:00-03:00) to early-morning (04:00-11:00) and evening (17:00-23:00) hours. However, LF Power had a significant difference between the groups during the afternoon hours (13:00-16:00). In contrast to ACCF/AHA and ASE/EACVI, RMSSD showed a high difference between values for HRVm2EF and other groups during the time periods 01:00-04:00 and 22:00-00:00. In addition, Alpha1 exhibited significant differences between HRVm2EF and other groups during the same aforementioned time periods. Additional features were also represented in the figure such as HF Norm, HF Peak, SD2, and PAS. Among these feature, SD2 have showed significant differences during the afternoon (12:00) along with the PAS (12:00-13:00 and 15:00-16:00).

Table 1. HRV features values (Mean±STD) at the most significantly different hours along with the *p*-value for one-way ANOVA test. The table shows features based on the ACCF/AHA and ASE/EACVI guidelines along with HRVEF/Jenks categories.

HRV Features		RMSSD		HF Peak		HF Power		LF Power		SD2		Alpha1		PAS	
Significant Hours		03:00-04:00	18:00-19:00	03:00-04:00	18:00-19:00	03:00-04:00	18:00-19:00	03:00-04:00	18:00-19:00	03:00-04:00	18:00-19:00	03:00-04:00	18:00-19:00	03:00-04:00	18:00-19:00
ACCF/AHA															
Mean ±	<i>pEF</i>	0.03±0.01 [‡]	0.02±0.01 ^{‡‡}	0.26±0.07	0.25±0.05	154.25±143.27 ^{‡‡}	140.34±112.04 ^{‡‡}	565.40±541.06	511.66±374.38 ^{‡‡}	0.11±0.04	0.11±0.04 ^{**}	1.37±0.16 ^{‡‡}	1.38±0.17 [*]	9.78±11.28	8.50±8.84
	<i>mEF</i>	0.02±0.02 ^{‡‡}	0.02±0.01 ^{‡‡}	0.27±0.06	0.28±0.08 ^{‡‡}	193.48±327.17 [‡]	128.56±162.71 ^{‡‡}	390.90±393.10 ^{‡‡}	302.63±305.87 [‡]	0.09±0.04	0.09±0.04 ^{**}	1.20±0.23 [*]	1.19±0.25 [*]	11.16±6.88	12.74±9.32
STD	<i>rEF</i>	0.04±0.04 ^{‡‡‡}	0.05±0.03 ^{‡‡‡}	0.25±0.06	0.21±0.07 ^{‡‡}	586.81±1003.10 ^{‡‡‡}	601.83±718.28 ^{‡‡‡}	1260.70±2206.90 ^{‡‡}	1034.10±1380.70 ^{‡‡‡}	0.11±0.06	0.11±0.05	1.20±0.20 ^{‡‡}	1.21±0.12	16.06±7.02	11.18±7.47
	ANOVA <i>p</i> -value	0.017	8.293x10⁻⁵	0.922	0.039	0.004	4.010x10⁻⁶	0.034	0.005	0.228	0.060	3.741x10⁻⁴	4.680x10⁻⁴	0.258	0.166
ASE/EACVI															
Mean ±	<i>pEF</i>	0.03±0.01 ^{‡‡}	0.02±0.01	0.27±0.07	0.25±0.06	151.19±150.00 ^{‡‡}	143.16±99.96 ^{‡‡}	590.07±594.88	533.95±372.42	0.10±0.04	0.11±0.04	1.38±0.16 [‡]	1.36±0.18 [‡]	9.43±11.90	8.80±9.55
	<i>mEF</i>	0.02±0.01 ^{‡‡}	0.02±0.01 ^{‡‡}	0.26±0.06	0.26±0.07	140.22±125.11	111.27±136.44 ^{‡‡}	392.06±263.53	364.32±366.61	0.10±0.05	0.12±0.05	1.32±0.15 ^{‡‡}	1.38±0.18 [‡]	10.08±7.84	9.80±8.37
STD	<i>rEF</i>	0.04±0.03 ^{‡‡‡}	0.03±0.03 ^{‡‡}	0.28±0.07	0.24±0.08	389.33±684.46 ^{‡‡}	331.55±503.13 ^{‡‡‡}	790.24±1421.00	602.14±931.27	0.11±0.05	0.10±0.05	1.17±0.23 ^{‡‡‡}	1.18±0.23 ^{‡‡}	14.28±7.27	11.45±7.78
	ANOVA <i>p</i> -value	0.019	0.020	0.767	0.708	0.024	0.010	0.301	0.355	0.926	0.338	9.490x10⁻⁵	7.180x10⁻⁴	0.186	0.519
HRVEF/Jenks															
Mean ±	<i>pEF</i>	0.03±0.01 ^x	0.02±0.01 [*]	0.23±0.07 ^{**}	0.25±0.06	163.12±198.28 [‡]	137.67±124.28 [‡]	525.93±514.43 ^{**}	464.38±368.92 ^{**}	0.10±0.04	0.11±0.04 ^{**}	1.36±0.16 ^{**‡}	1.25±0.23	8.47±6.36 ^x	8.36±8.84 ^{‡‡}
	<i>mEF</i>	0.04±0.02 ^o	0.06±0.07 [‡]	0.29±0.07 ^{**}	NA	1906.80±1498.70 [‡]	1175.40±936.91 [‡]	3627.50±4211.30 ^{**‡‡}	2345.20±2545.90 ^{**‡‡}	0.12±0.04	0.19±0.06 ^{**‡‡}	0.99±0.38 [‡]	0.94±0.45 ^o	8.92±7.57 ^o	NA
STD	<i>m2EF</i>	0.14±0.01 ^{xxo+}	NA	NA	NA	NA	NA	NA	NA	0.11±0.10	NA	NA	1.38±0.16 ^{oo+}	40.45±30.11 ^{xxo+}	NA
	<i>rEF</i>	0.02±0.01 ⁺	0.03±0.01 [‡]	0.27±0.06	0.21±0.07	146.84±174.36 ^{‡‡}	257.67±262.94 ^{‡‡}	471.80±532.76 ^{‡‡}	597.08±675.13 ^{‡‡}	0.06±0.02	0.10±0.04 ^{‡‡}	1.18±0.20 [‡]	1.19±0.20 ⁺	16.50±6.70 ⁺	12.46±8.59 ^{‡‡}
ANOVA <i>p</i> -value	1.971x10⁻¹⁰	0.017	0.027	0.120	1.275x10⁻¹⁴	1.201x10⁻¹³	4.978x10⁻⁸	1.765x10⁻⁶	0.018	3.800x10⁻⁴	2.860x10⁻⁵	3.102x10⁻⁵	6.469x10⁸	0.042	

[‡]: Significant difference between *pEF* and *mEF* groups (p-value < 0.01), ^{**}: Significant difference between *pEF* and *mEF* groups (p-value < 0.05).
^{‡‡}: Significant difference between *rEF* and *mEF* groups (p-value < 0.01), ^{‡‡‡}: Significant difference between *rEF* and *mEF* groups (p-value < 0.05).
[‡]: Significant difference between *pEF* and *rEF* groups (p-value < 0.01), ^{‡‡}: Significant difference between *pEF* and *rEF* groups (p-value < 0.05).
^x: Significant difference between *pEF* and *m2EF* groups (p-value < 0.01), ^{xx}: Significant difference between *pEF* and *m2EF* groups (p-value < 0.05).
^o: Significant difference between *mEF* and *m2EF* groups (p-value < 0.01), ^{oo}: Significant difference between *mEF* and *m2EF* groups (p-value < 0.05).
⁺: Significant difference between *rEF* and *m2EF* groups (p-value < 0.01), ⁺⁺: Significant difference between *rEF* and *m2EF* groups (p-value < 0.05).
HRVm2EF: HRV with mid-range 2 EF group (second stage mid-range category).
NA: Not Applicable (only 3 HRVEF categories were determined).

Discussion

This study demonstrated a strong association between LVEF and HRV during specific time periods in the circadian cardiac rhythm. The new index, HRVEF, allowed for multi-time estimations of LVEF throughout the day and night based on the functionality of the heart as represented by HRV features. The results suggest that HRV could be a promising automated, economical, and accurate alternative to the once-off echocardiography test for LVEF measurements and heart failure prognosis.

Although the Jenks algorithm is most commonly used in choropleth maps clustering, it is interesting to observe its ability in segmenting data for medical applications. Jenks natural breaks and HRV allowed for an expanded view on the distribution of LVEF among CAD patients. Unlike the current LVEF guidelines where a patient is classified based on their echocardiography LVEF level, the Jenks breaks optimization for certain HRV features helped in assigning patients into LVEF groups on an hourly basis. In addition, the optimization was not only based on a three groups scenario, but also indicated that for certain HRV features and time intervals the patients could be divided into two and four LVEF groups.

The variations in the number of group scenarios across the 24 hours show that a single feature may be observed very similarly across patients with different LVEF levels, thus, resulting in fewer LVEF groups. Taking HF Peak feature as an example, which is a good indicator for the cardiac parasympathetic nerve activity²², a two groups scenario of HRVrEF and HRVpEF was suggested during the evening time periods (18:00-21:00). Therefore, at these hours, the need for an additional mid-range group was not required. This indicates that there is a greater effect of parasympathetic system modulation of cardiac rhythm in the evening hours and hence a greater difference between postulated groups with less intragroup variance.

However, at other time periods, a wider distribution of patients based on LVEF was observed, leading to higher number of LVEF groups (three or four). An example of this was the LF Power results, that mainly represents the sympathetic branch of the ANS²², where a three groups scenario was suggested with a ≥ 0.70 GVF during the late-night to early-morning (01:00-08:00) and evening (18:00-22:00) hours. Furthermore, features such as SD2 which reflects sympathetic activation²³ and PAS that mainly represents changes in the heart rate acceleration²⁴ suggested an additional HRVm2EF group due to a greater variability in these HRV features between patients during the early-morning time periods. This indicates that the early morning sympathetic surge affects patients with similar LVEF differently due to some difference in heart mechanics or rhythm generator which is identified by the more sensitive HRV feature.

The analysis of HRV provides information about the overall status of ANS modulation of cardiac rhythm. Although the use of HRV features for estimating LVEF is effective when investigating the effect of the ANS on cardiac rhythm, the relationship between HRV indices and LVEF has not been well defined in CAD patients in the literature. One of the reasons could be that HRV analysis is affected by the presence of cardiac arrhythmias and oscillatory modulations due to sources other than intrinsic cardiac regulation (e.g. SA node)²⁵. In the current study, pre-processing of the ECG ensured removal of heart rate artifacts and led to significant findings. High-frequency HRV features (HF Peak, HF Power), SD2, and Alpha1, which are associated with parasympathetic nervous system activities correlated with the degree of LVEF. This was especially the case during the late-night to early-morning (00:00-06:00) and evening (19:00-23:00) time periods, which are high-risk times for cardiac arrest. These findings match the previous observations reported in the literature that suggested time intervals such as 6 pm to 12 am as times of higher mortality due to cardiac infarction²⁶. The current findings could open up the possibility of cardiovascular chronopharmacology where HRV provides additional information on the association with various degrees of LVEF during specific hourly (diurnal and nocturnal) intervals.

Relative to the current ACCF/AHA and ASE/EACVI guidelines for heart failure classification based on LVEF, results for HRVEF indicated some changes in the range of accepted heart failure classes based on LVEF classifications. Features such as RMSSD and pNN50 suggested a range for HRVmEF group between $35\% \geq EF \geq 44\%$. This agrees to some extent with the ACCF/AHA guidelines for patients with more than 40% LVEF, however, based on the ACCF/AHA guidelines, few HFREF patients were assigned to the the mid-range group. On the other hand, more patients were in the mid-range category based on ASE/EACVI but these guidelines recommend a range of 50% to 55% for mid-range LVEF, which is significantly different to the HRVEF estimations shown herein. The aforementioned findings show patients with $EF \geq 44\%$ at normal heart functionality during the late-night to early-morning and evening time periods. However, these patients become at-risk of developing cardiac abnormalities leading to heart failure by moving from being HRVpEF to HRVmEF or HRVrEF during the afternoon as seen in the full spectrum of the RMSSD and pNN50 features (Supplementary Figure S2). Both of these features reflect the modulation of the parasympathetic activity and their variations are usually linked to heart failure. For example, lower RMSSD has been shown to be associate with an increase in all-cause mortality²⁷. Furthermore, RMSSD variations are mainly due to the fast frequency modulation of the SA-node via the parasympathetic input at about 1-4 beats coupled with a slower sympathetic input between 4-10 beats. Additionally to these features, HF Power and SD1 showed similar results suggesting less activation in the parasympathetic branch at the aforementioned time periods.

In contrast to the ACCF/AHA and ASE/EACVI guidelines where the upper cut-off value for the mid-range group was set at 50 and 55, respectively, a wider distribution for the HRVmEF group was observed with the SD2 feature during the early-morning (05:00-08:00) time periods. Furthermore, an additional HRVm2EF group was suggested due to the higher

214 variability between patients at higher LVEF levels ($EF \geq 70$). The additional group with a LVEF range between 73 to 75
215 suggests greater force of the heart activity for patients at those levels during early-morning hours and distinguished them from
216 the preserved and mid-range LVEF groups. However, most of the HRVmEF patients ($EF \geq 44$) were considered earlier as
217 preserved patients during the late-night (01:00-04:00) hours, which matches the range for the ACCF/AHA guidelines of $EF \geq$
218 50, before they start dividing up in the early-morning. This shows that these patients were possibly not at-risk of a cardiac event
219 associate with heart failure during these time periods as seen from the SD2 feature that reflects sympathetic branch activation,
220 as sympathetic activation usually implies greater risk of a cardiac event. Looking at the overall 24-hour spectrum of this feature,
221 patients started to divide into two groups with the majority being considered as HRVrEF patients before they start returning
222 back during the late-night hours. Knowledge about the at-risk time periods allows for optimal medications to prevent any
223 further development of a cardiac event in the future. Although the PAS HRV feature showed no clear grouping of patients
224 throughout the 24-hour cycle, it had a similar pattern during the early-morning time periods as seen in Supplementary Figure
225 S2, which supports the claim that the early morning is a high-risk time periods for cardiac arrest associated with heart failure.

226 Comparing the significant differences across LVEF groups based on ACCF/AHA, ASE/EACVI, and HRVEF, the statistical
227 discrimination ability was higher between the HRVEF groups as seen in Table 1. The mid-range group (blue) had a wider
228 difference in values compared to other groups for features such as RMSSD, HF Power, and LF Power, which was not clearly
229 shown in the observations associated with the ACCF/AHA and ASE/EACVI guidelines. In the proposed HRVEF guidelines,
230 greater differences in values were observed for the reduced patients, whereas the patients with rEF had closer values to the
231 preserved group. However, patients in the rEF with very low LVEF values as discussed earlier showed severe symptoms of
232 heart failure that distinguished them from the preserved patients. The more discrimination ability of mid-range patients, the less
233 ambiguity there is in the classification of such patients and the better the treatment and medication process. In addition,
234 the newly introduced HRVm2EF group was found to be significantly different at certain time periods, in contrast to the ACCF/AHA
235 an ASE/EACVI guidelines for LVEF. This may indicate additional information about high functionality of the heart at specific
236 time periods for specific patients²⁸.

237 **Materials and Methods**

238 **Data set and Patients Enrollment**

239 The selected CAD patients data set was obtained from the Intercity Digital ECG Alliance (IDEAL) study of the University of
240 Rochester Medical Center Telemetric and Holter ECG Warehouse (THEW) archives²⁹. Informed consent was obtained from
241 all participants. The database measurements were conducted according to Title 45, U.S. Code of Federal Regulations, Part
242 46, Protection of Human Subjects (Revised November 13, 2001 - effective December 13, 2001) and in accordance with the
243 Declaration of Helsinki. Furthermore, the Research Subject Review Board of the University of Rochester approved the IDEAL
244 protocol³⁰.

245 The original data set included 271 patients that are in a stable condition after 2+ months since their latest cardiac event.
246 All patients passed the eligibility criteria that included having exercise-induced ischemia, a record of myocardial infarction
247 (MI), at least one vessel narrowing $>75\%$, stable ischemic heart disease, and sinus rhythm. Furthermore, these patients did not
248 suffer from dilated cardiomyopathy, unstable angina, congestive heart failure (CHF), previous coronary artery bypass surgery
249 (CABG), and cerebral/renal vascular diseases.

250 All patients had a 24-hour Holter ECG recorded using a three pseudo-orthogonal lead configuration and an echocardiography
251 test to measure their LVEF level. In the current study, patients with hypertension (110), diabetes (9), or both (27) were excluded
252 from any further analysis as these comorbidities have a strong effect on HRV. This resulted in including 92 patients in the
253 finalized data set after additionally removing patients with missing ECG or HRV annotations. The complete demographic
254 information of patients within each LVEF group following both LVEF guidelines is provided in Table 2. All Holter ECG
255 recordings were pre-processed to remove noise and ectopic beats using the signal-dependent rank order mean (SD-ROM) and
256 adaptive filtering techniques to ensure no abnormalities in the data.

257 **Circadian HRV features**

258 The database does not mention the starting time of each recording, therefore the starting point for analysis of the 24-hour
259 circadian rhythm was fixed prior to analysis. Cosinor fitting analysis^{12,31} was performed to ensure that all HRV data started
260 from 12:00AM. In this analysis, a cosine function is used to fit each HRV data by calculating the midline estimating statistic of
261 rhythm (MESOR-M), amplitude (Amp), and acrophase (AC). The reference angle was chosen to be 0° , referring to 12:00AM,
262 with an increase of 15° for every hour obtained after converting the 24-hour AC into angle data using $(360/24)$. For example, if
263 an AC of 65.30° was obtained for a patients' HRV data, it is converted to the corresponding hour of 04:00AM.

264 After obtaining the time-fixed data set for all patients, the hourly HRV features were extracted from time- and frequency-
265 domain metrics³², non-linear metrics including Poincare plot, detrended fluctuation analysis (DFA), and multi-scale entropy

Table 2. Demographic information of patients within each LVEF group according to the ACCF/AHA and ASE/EACVI gold-standard guidelines.

LVEF Guidelines	Patients (n)	Sex (M/F)	Age (Mean±Std) (yrs)	BMI (Mean±Std) (kg/m ²)	BPS (Mean±Std) (mmHg)	BPD (Mean±Std) (mmHg)	LVDS (Mean±Std) (mm)	LVDD (Mean±Std) (mm)	Smoking (Yes/No)	Syncope (n)	VT (n)	
ACCF/AHA	<i>pEF</i>	65	57/8	35-84 (56.42±11.27)	17.99-36.33 (25.83±3.77)	100-160 (121.77±12.51)	50-90 (76.39±7.29)	22-55 (33.87±7.53)	33-68 (51.02±6.49)	50/15	7	5
	<i>mEF</i>	19	19/0	40-80 (59.22±11.42)	22.55-31.26 (27.13±2.51)	90-150 (123.47±15.12)	60-95 (75.21±10.07)	31-57 (42.14±7.52)	37-71 (54.72±10.91)	14/5	1	2
	<i>rEF</i>	8	8/0	41-70 (54.25±9.74)	21.55-36.85 (26.75±5.25)	90-130 (116.25±13.30)	60-80 (70.13±6.11)	33-66 (46.21±12.73)	28-75 (55.71±16.32)	7/1	1	4
ASE/EACVI	<i>pEF</i>	51	44/7	35-79 (55.98±11.11)	19.72-36.33 (25.56±3.67)	100-160 (121.75±12.96)	60-90 (75.53±6.44)	22-55 (33.37±6.78)	33-59 (50.53±4.98)	40/11	3	6
	<i>mEF</i>	20	19/1	40-84 (58.68±11.57)	17.99-33.96 (26.66±3.71)	100-140 (122.30±11.13)	50-90 (78.15±9.72)	26-45 (37.23±4.48)	40-68 (54.08±7.03)	13/7	5	0
	<i>rEF</i>	21	21/0	40-80 (57.00±11.15)	21.55-36.85 (27.21±3.63)	90-150 (120.76±15.62)	60-95 (73.33±9.13)	31-66 (44.56±9.03)	28-75 (55.02±11.79)	19/2	1	5
Overall (per-guideline)	92	84/8	35-84 (56.79±11.14)	17.99-36.85 (26.18±3.70)	90-160 (121.64±13.12)	50-95 (75.60±7.96)	22-66 (36.80±9.24)	28-75 (52.21±8.84)	71/21	9	11	

BMI: Body mass index (kg/m³), BPS: Systolic blood pressure (mmHg), BPD: Diastolic blood pressure (mmHg), LVDS: Left ventricular end-systolic dimension (mm), LVDD: Left ventricular end-diastolic dimension (mm), VT: Ventricular tachycardia

(MSE)^{33,34}. Furthermore, the newly introduced fragmentation metrics³⁵ were also included. All features used in the current study are listed below and briefly defined in Supplementary Table S1:

- **Time-domain:** AVNN, SDNN, RMSSD, pNN50, SEM
- **Frequency-domain:** BETA, HF Norm, HF Peak, HF Power, LF Norm, LF Peak, LF Power, LF/HF, Total Power, VLF Norm, VLF Power
- **Non-linear:** SD1, SD2, Alpha1, Alpha2, Sample Entropy
- **Fragmentation:** PIP, IALS, PSS, PAS

Missing Data imputation

Moving mean

In this basic data imputation approach, each missing value for each feature is replaced by a moving mean value with a pre-defined window length. In this work, the window size is determined to be 5.

K-nearest neighbour (kNN)

In kNN, missing values are filled based on the information taken from other available values obtained from neighbours³⁶. If the targeted missing value is expected to be numerical, which is the case in current data, the missing value is filled with a value obtained from the k nearest neighbours. In this work, the number of k neighbours was selected to be 5.

Multiple imputation by chained equation (MICE)

The MICE algorithm suggests the use of multiple imputations in addressing missing values as opposed to single imputation methods. In multiple imputations, missing values are imputed multiple times iteratively with values drawn from a posterior predictive distribution of the missing values conditional on the available values³⁷. An advantage of the MICE algorithm is the ability to impute data on a variable-by-variable basis³⁸.

Neural networks

The selected neural network architecture is based on Datawig framework³⁹, which provides deep learning models to impute missing values. In this approach, features are learned through a symbolic API of Apache MXNET. The advantage of utilizing this technique is the ability to learn from all other features prior to any prediction of missing features values. In this work, each missing value is initially replaced by the average value of its variable across patients. The encoder and featurizer were selected to be numerical corresponding to the type of data imputed.

292 **Random forest (RF)**

293 RF is machine learning algorithm that utilizes ensembles of decision trees through bootstrap aggregation to make decisions⁴⁰.
294 In the context of data imputation, all missing value were initially replaced by their corresponding variable average. The
295 training/prediction of missing values was performed iteratively as a regression problem, where on each iteration, missing values
296 are updated based on learned features. The imputation process is performed as a prediction provided after learning from all
297 other features. In this work, the total number of decision trees was specified to be 50.

298 **Jenks natural breaks optimization**

299 Jenks natural breaks is a 1D clustering technique used to arrange data values into different classes based on within-classes
300 variance calculations. The natural ranges in a dataset of values refer to the most optimal class ranges found naturally in the data
301 that form natural groups of values with similar characteristics⁴¹. To optimize natural classes, the algorithm seeks minimizing
302 the deviation of values in each class from the class mean. In addition, it tries to maximize the deviation of values from the mean
303 of other classes. In other words, it reduces the within-class variance and increases the between-classes variance^{41,42}.

304 This technique performs three major steps iteratively to optimize natural breaks. These steps are:

305 1) Find the sum of squared deviations from the current array mean (SDAM) as follows,

$$SDAM = \sum_{i=1}^n (x_i - \bar{x})^2, \quad (1)$$

306 where i denotes the array value, n is the total number of values in the array, and \bar{x} is the mean of values within the array.

307 2) Find the sum of squared deviations from each class mean (SDCM) by,

$$SDCM = \sum_{k=1}^m \sum_{i=1}^n (x_{ik} - \bar{x}_k)^2, \quad (2)$$

308 where k is the class number, m is the total number of classes, x_{ik} is the i^{th} value in the class array, and \bar{x}_k is the mean of values
309 within the class.

310 3) Decide if a move of one unit from the class with a larger SDCM to an adjacent class with a lower SDCM is needed. This
311 step is needed to ensure that the within-class deviations are minimized.

312 A goodness of variance fit (GVF) parameter is measured at the end of the clustering process to evaluate the variance fitting
313 performance. GVF ranges between 1 (perfect fit) and 0 (worst fit) and is calculated as follows:

$$GVF = \frac{SDAM - SDCM}{SDAM}. \quad (3)$$

314 To ensure the best possible clustering process of the data, the GVF value was obtained when for three clustering scenarios;
315 namely 2, 3, and 4 classes. The best class clustering scenario was selected at each hour and for every HRV feature once the
316 GVF reaches a value of more than or equal to 0.70. In the case of 2-class, HRVEF groups were named as HRV with reduced EF
317 (HRVrEF) and HRV with preserved EF (HRVpEF). For 3-class scenario, an additional group was included between the two
318 aforementioned groups as HRV with mid-range EF (HRVmEF). Finally, the 4-class scenario included an additional HRVmEF
319 named as HRV with mid-range 2 EF (HRVm2EF).

320 **Statistical analysis of features**

321 To evaluate HRV features on an hourly basis, the average feature value was calculated for each of ACCF/AHA and ASE/EACVI
322 LVEF classification guidelines as well as the HRVEF groups obtained by the Jenks algorithm. An one-way analysis of variance
323 (ANOVA) test was carried out to compare the mean values of each feature across LVEF groups with significant difference
324 between groups at ($p < 0.05$)⁴³. Furthermore, a multiple comparison post-hoc analysis (Tukey *post-hoc* test)⁴⁴ was applied.

325 **References**

- 326 1. G. Ramani and P. Uber and M. Mehra. Chronic heart failure: contemporary diagnosis and management. In *Mayo Clinic*
327 *Proceedings*, vol. 85, 180–195 (Elsevier, 2010).
- 328 2. P. Ponikowski and S. Anker and K. AlHabib and M. Cowie and T. Force and S. Hu and T. Jaarsma and H. Krum and V.
329 Rastogi and L. Rohde and others. Heart failure: preventing disease and death worldwide. *ESC Heart Failure* **1**, 4–25
330 (2014).

- 331 3. J. Rodriguez and A. Voss and P. Caminal and A. Bayés-Genis and B. Giraldo. Characterization and classification of patients
332 with different levels of cardiac death risk by using Poincaré plot analysis. In *2017 39th Annual International Conference of*
333 *the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1332–1335 (IEEE, 2017).
- 334 4. N. El-Sherif and A. Khan and J. Savarese and G. Turitto. Pathophysiology, risk stratification, and management of sudden
335 cardiac death in coronary artery disease. *Cardiology Journal* **17**, 4–10 (2010).
- 336 5. C. Yancy and M. Jessup and B. Bozkurt and J. Butler and D. Casey and M. Drazner and G. Fonarow and S. Geraci and T.
337 Horwich and J. Januzzi and others. 2013 ACCF/AHA guideline for the management of heart failure: executive summary: a
338 report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines.
339 *Journal of the American College of Cardiology* **62**, 1495–1539 (2013).
- 340 6. R. Lang and M. Bierig and R. Devereux and F. Flachskampf and E. Foster and P. Pellikka and others. Recommendations
341 for chamber quantification. *European journal of echocardiography* **7**, 79–108 (2006).
- 342 7. J. Januzzi and Y. Chandrashekhar. Strain echocardiography: the new gold standard for imaging ventricular function?
343 (2017).
- 344 8. T. Foley and S. Mankad and N. Anavekar and C. Bonnichsen and M. Morris and T. Miller and P. Araoz. Measuring left
345 ventricular ejection fraction—techniques and potential pitfalls. *European Cardiology* **8**, 108–114 (2012).
- 346 9. Khandoker, Ahsan H and Jelinek, Herbert F and Moritani, Toshio and Palaniswami, Marimuthu. Association of cardiac au-
347 tonomic neuropathy with alteration of sympatho-vagal balance through heart rate variability analysis. *Medical engineering*
348 *@AND@ physics* **32**, 161–167 (2010).
- 349 10. Shaffer, F. & Ginsberg, J. An overview of heart rate variability metrics and norms. *Frontiers in public health* **5**, 258 (2017).
- 350 11. H. Huikuri and P. Stein. Heart rate variability in risk stratification of cardiac patients. *Progress in cardiovascular diseases*
351 **56**, 153–159 (2013).
- 352 12. AlKhadari, M., Jelinek, H., Werghe, N., Hadjileontiadis, L. & Khandoker, A. Estimating left ventricle ejection fraction
353 levels using circadian heart rate variability features and support vector regression models. *IEEE Journal of Biomedical and*
354 *Health Informatics* (2020).
- 355 13. AlKhadari, M., Jelinek, H., Werghe, N., Hadjileontiadis, L. & Khandoker, A. Discrimination amongst various degrees of
356 left ventricular ejection fraction in cad patients using circadian heart rate variability features. In *2020 11th Conference of*
357 *the European Study Group on Cardiovascular Oscillations (ESGCO)*, 1–2 (IEEE, 2020).
- 358 14. AlKhadari, M., Jelinek, H. F., Werghe, N., Hadjileontiadis, L. J. & Khandoker, A. H. Investigating circadian heart rate
359 variability in coronary artery disease patients with various degrees of left ventricle ejection fraction. In *2020 42nd Annual*
360 *International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 714–717 (IEEE, 2020).
- 361 15. Dekker, J. M. *et al.* Low heart rate variability in a 2-minute rhythm strip predicts risk of coronary heart disease and
362 mortality from several causes: the aric study. *Circulation* **102**, 1239–1244 (2000).
- 363 16. Kiyono, K., Hayano, J., Watanabe, E., Struzik, Z. R. & Yamamoto, Y. Non-gaussian heart rate as an independent predictor
364 of mortality in patients with chronic heart failure. *Heart Rhythm* **5**, 261–268 (2008).
- 365 17. Stein, P. K. *et al.* Novel measures of heart rate variability predict cardiovascular mortality in older adults independent of
366 traditional cardiovascular risk factors: the cardiovascular health study (chs). *Journal of Cardiovascular Electrophysiology*
367 **19**, 1169–1174 (2008).
- 368 18. Huikuri, H. V. *et al.* Prediction of fatal or near-fatal cardiac arrhythmia events in patients with depressed left ventricular
369 function after an acute myocardial infarction. *European heart journal* **30**, 689–698 (2009).
- 370 19. Hayano, J. *et al.* Increased non-gaussianity of heart rate variability predicts cardiac mortality after an acute myocardial
371 infarction. *Frontiers in Physiology* **2**, 65 (2011).
- 372 20. AlKhadari, M., Islayem, D. K., Alskafi, F. A. & Khandoker, A. H. Predicting hypertensive patients with higher risk of
373 developing vascular events using heart rate variability and machine learning. *IEEE Access* **8**, 192727–192739 (2020).
- 374 21. Hayano, J. *et al.* Survival predictors of heart rate variability after myocardial infarction with and without low left ventricular
375 ejection fraction. *Frontiers in Neuroscience* **15**, 32 (2021).
- 376 22. Billman, G. E. The lf/hf ratio does not accurately measure cardiac sympatho-vagal balance. *Frontiers in Physiology* **4**, 26
377 (2013).
- 378 23. Brennan, M., Palaniswami, M. & Kamen, P. Do existing measures of poincare plot geometry reflect nonlinear features of
379 heart rate variability? *IEEE Transactions on Biomedical Engineering* **48**, 1342–1347 (2001).

- 380 **24.** Costa, M. D., Davis, R. B. & Goldberger, A. L. Heart rate fragmentation: a symbolic dynamical approach. *Frontiers in*
381 *Physiology* **8**, 827 (2017).
- 382 **25.** Tundo, F. *et al.* Heart rate turbulence and left ventricular ejection fraction in chagas disease. *EP Europace* **7**, 197–203
383 (2005).
- 384 **26.** Schwartz, Bryan G and Mayeda, Guy S and Burstein, Steven and Economides, Christina and Kloner, Robert A. When and
385 why do heart attacks occur? Cardiovascular triggers and their potential role. *Hospital Practice* **38**, 144–152 (2010).
- 386 **27.** Drawz, P. E. *et al.* Heart rate variability is a predictor of mortality in chronic kidney disease: a report from the cric study.
387 *American Journal of Nephrology* **38**, 517–528 (2013).
- 388 **28.** Hägglund, Harriet and Uusitalo, Arja and Peltonen, Juha E and Koponen, Anne S and Aho, Jyrki and Tiinanen, Suvi and
389 Seppänen, Tapio and Tulppo, Mikko and Tikkanen, Heikki Olavi. Cardiovascular autonomic nervous system function and
390 aerobic capacity in type 1 diabetes. *Frontiers in physiology* **3**, 356 (2012).
- 391 **29.** University of Rochester Medical Center. Telemetric and holter ECG warehouse.
- 392 **30.** Burattini, L. & Burattini, R. Characterization of repolarization alternans in the coronary artery disease. In *Coronary Artery*
393 *Diseases*, 91–112 (IntechOpen, 2012).
- 394 **31.** Jelinek, H. F. *et al.* Temporal dynamics of the circadian heart rate following low and high volume exercise training in
395 sedentary male subjects. *European journal of applied physiology* **115**, 2069–2080 (2015).
- 396 **32.** M. Malik. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use: Task force of the
397 European Society of Cardiology and the North American Society for Pacing and Electrophysiology. *Annals Noninvasive*
398 *Electrocardiol.* **1**, 151–181 (1996).
- 399 **33.** Peng, CHUNG-KANG and Hausdorff, JEFFREY M and Goldberger, AL. Fractal mechanisms in neuronal control: human
400 heartbeat and gait dynamics in health and disease. *Nonlinear Dyn. Self-organization Biomed.* 66–96 (2000).
- 401 **34.** Costa, Madalena and Goldberger, Ary L and Peng, C-K. Multiscale entropy analysis of biological signals. *Physical review*
402 *E* **71**, 021906 (2005).
- 403 **35.** Costa, Madalena D and Davis, Roger B and Goldberger, Ary L. Heart rate fragmentation: a new approach to the analysis
404 of cardiac interbeat interval dynamics. *Frontiers in Physiology* **8**, 255 (2017).
- 405 **36.** Zhang, S. Nearest neighbor selection for iteratively knn imputation. *J. Syst. Softw.* **85**, 2541–2552 (2012).
- 406 **37.** Buuren, S. v. & Groothuis-Oudshoorn, K. mice: Multivariate imputation by chained equations in r. *Journal of statistical*
407 *software* 1–68 (2010).
- 408 **38.** Bouhlila, D. S. & Sellaouti, F. Multiple imputation using chained equations for missing data in timss: a case study.
409 *Large-scale Assessments in Education* **1**, 4 (2013).
- 410 **39.** Biessmann, F. *et al.* Datawig: Missing value imputation for tables. *Journal of Machine Learning Research* **20**, 1–6 (2019).
- 411 **40.** Breiman, L. & Cutler, A. Random forests. department of statistics. *University of California, Berkeley.* <http://stat-www.berkeley.edu/users/breiman/Random-Forests> (2004).
- 412
- 413 **41.** Jenks, G. F. The data model concept in statistical mapping. *International Yearbook of Cartography* **7**, 186–190 (1967).
- 414 **42.** McMaster, R. In memoriam: George f. jenks (1916-1996). *Cartography and Geographic Information Systems* **24**, 56–59
415 (1997).
- 416 **43.** Hamada, M. & Wu, J. *Experiments: planning, analysis, and parameter design optimization* (Wiley New York, 2000).
- 417 **44.** Hochberg, Y. & Tamhane, A. Multiple comparison procedures. 1987. *John @AND@ Wiley Sons, New York* (1987).

418 **Acknowledgements**

419 The authors would like to acknowledge Khwaja Y. Hasan from the Cardiology Department at Cleveland Clinic, Abu Dhabi
420 for his advice on classification of LVEF in heart failure patients and the Telemetric and Holter ECG Warehouse (THEW) for
421 providing the data.

422 **Author contributions statement**

423 M.A., H.F.J., L.J.H, A.H.K. designed research idea. M.A. performed research. S.S. developed the de-noising approach. M.A.,
424 H.F.J., L.J.H, A.H.K. analyzed data and results. M.A. wrote the paper. H.F.J., L.J.H, A.H.K. edited the paper. All authors
425 reviewed the manuscript.

426 **Competing interests**

427 The authors declare no competing interests.

Figures

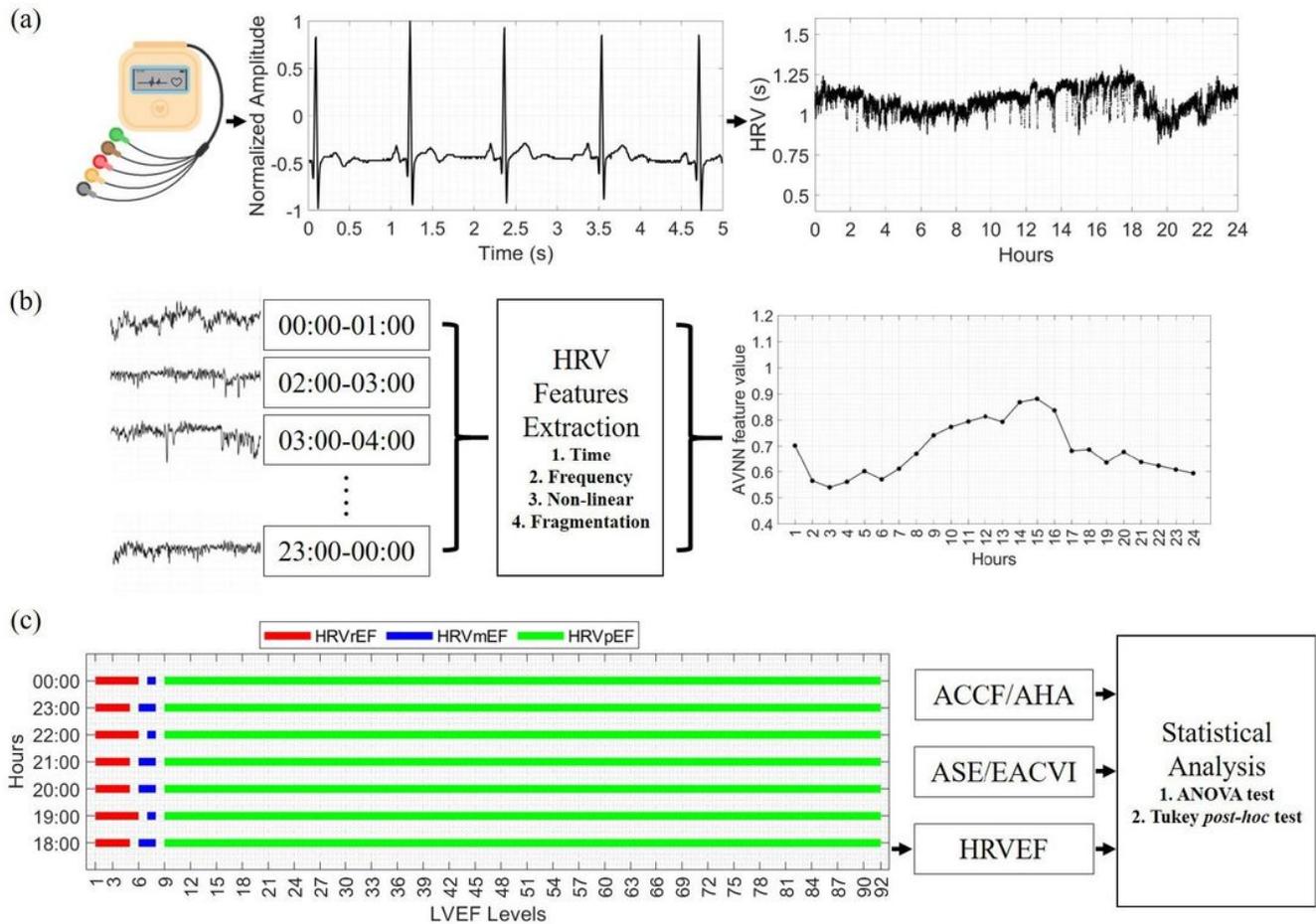


Figure 1

Flowchart for the generation of per-hour heart rate variability ejection fraction (HRVEF) index. (a) HRV acquisition from 24-hour Holter ECG recordings of coronary artery disease (CAD) patients. (b) Per-hour extraction of HRV features from time, frequency, non-linear, and fragmentation metrics (showing average normal-to-normal feature (AVNN)). (c) Utilizing Jenks natural breaks algorithm to categorize patients with respect to their left ventricular ejection fraction (LVEF) levels and statistically comparing them with gold standard guidelines.

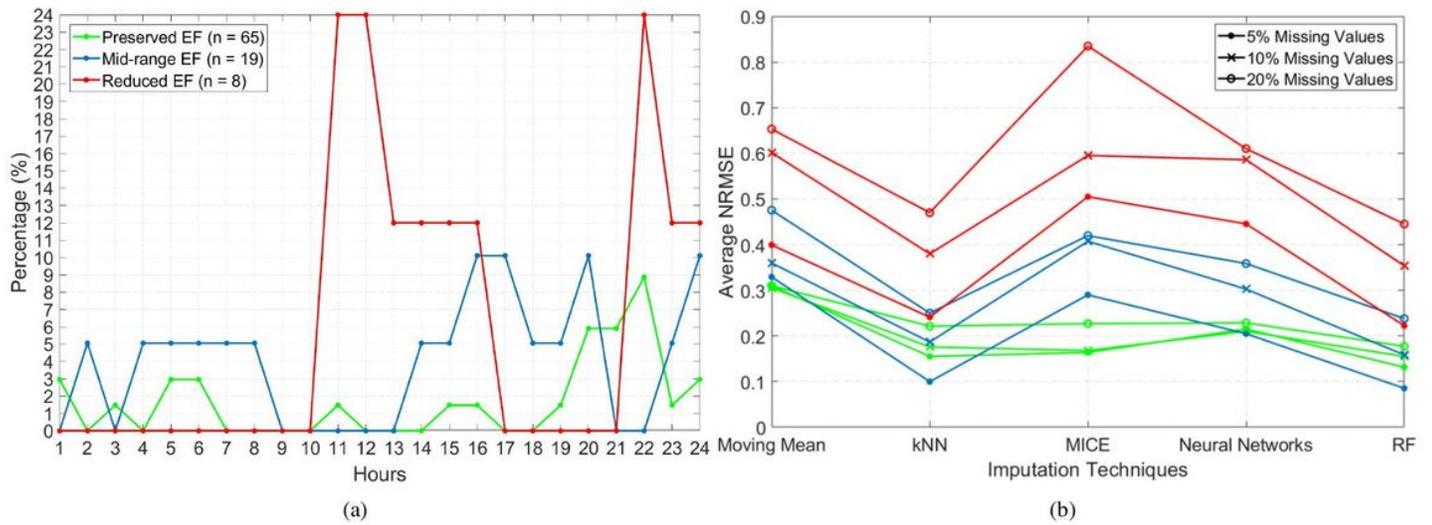


Figure 2

HRV feature imputation process to ensure a complete features dataset with no missing values. (a) Percentage of missing values across the 24-hour HRV data for each LVEF patient groups. (b) Average normalized root mean square error (NRMSE) for the five data imputation techniques. Three missing values scenarios (Dot: 5%, Cross: 10%, and Circle: 20%) were applied on the preserved (green), mid-range (blue), and reduced (red) LVEF groups to evaluate each technique.

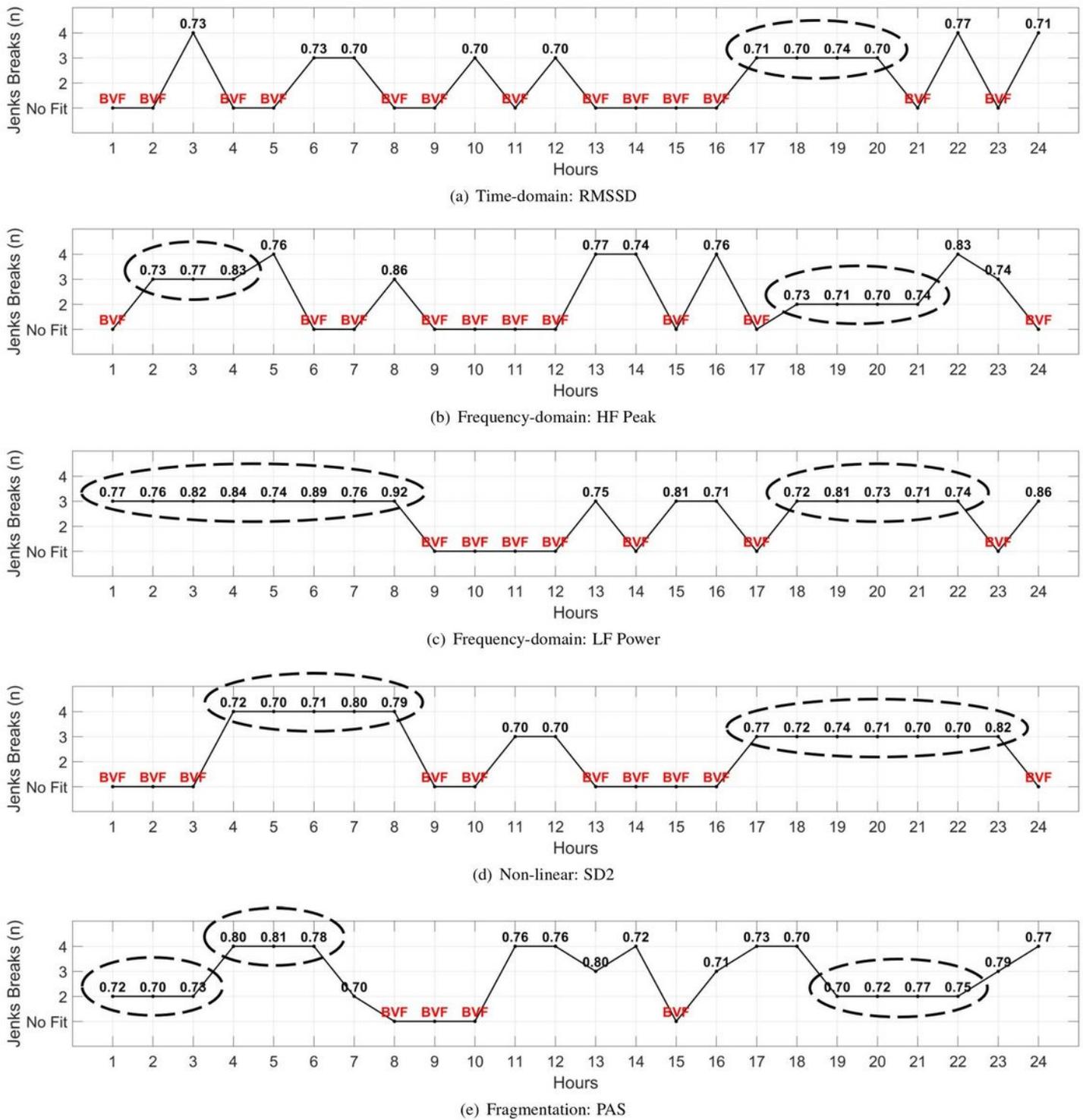


Figure 3

Heart rate variability ejection fraction (HRVEF) grouping using Jenks for selected HRV features. The figure shows the number of Jenks breaks whenever the goodness of variance fit (GVF) have reached ≥ 0.70 for (a) time-domain: RMSSD. (b) Frequency-domain: HF Peak. (c) Frequency-domain: LF Power. (d) Non-linear: SD2. (e) Fragmentation: PAS. Bad variance fit (BVF) was assigned for $GVF < 0.70$. The dashed circles denote having three or more consecutive ≥ 0.70 GVF

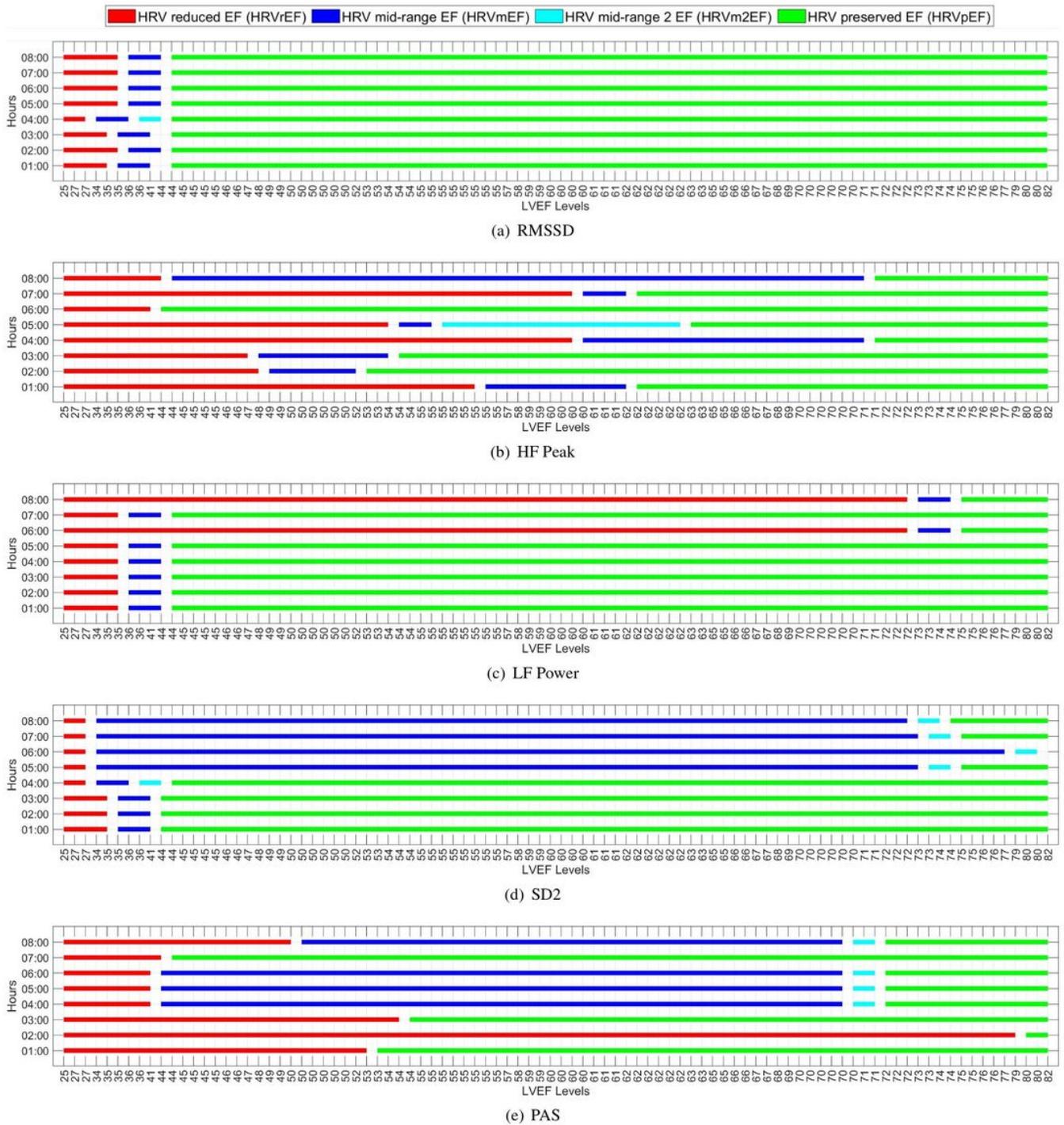


Figure 4

The late-night (01:00-03:00) to early-morning (04:00-08:00) hours distribution of heart rate variability ejection fraction (HRVEF) groups across patients based on Jenks natural breaks.

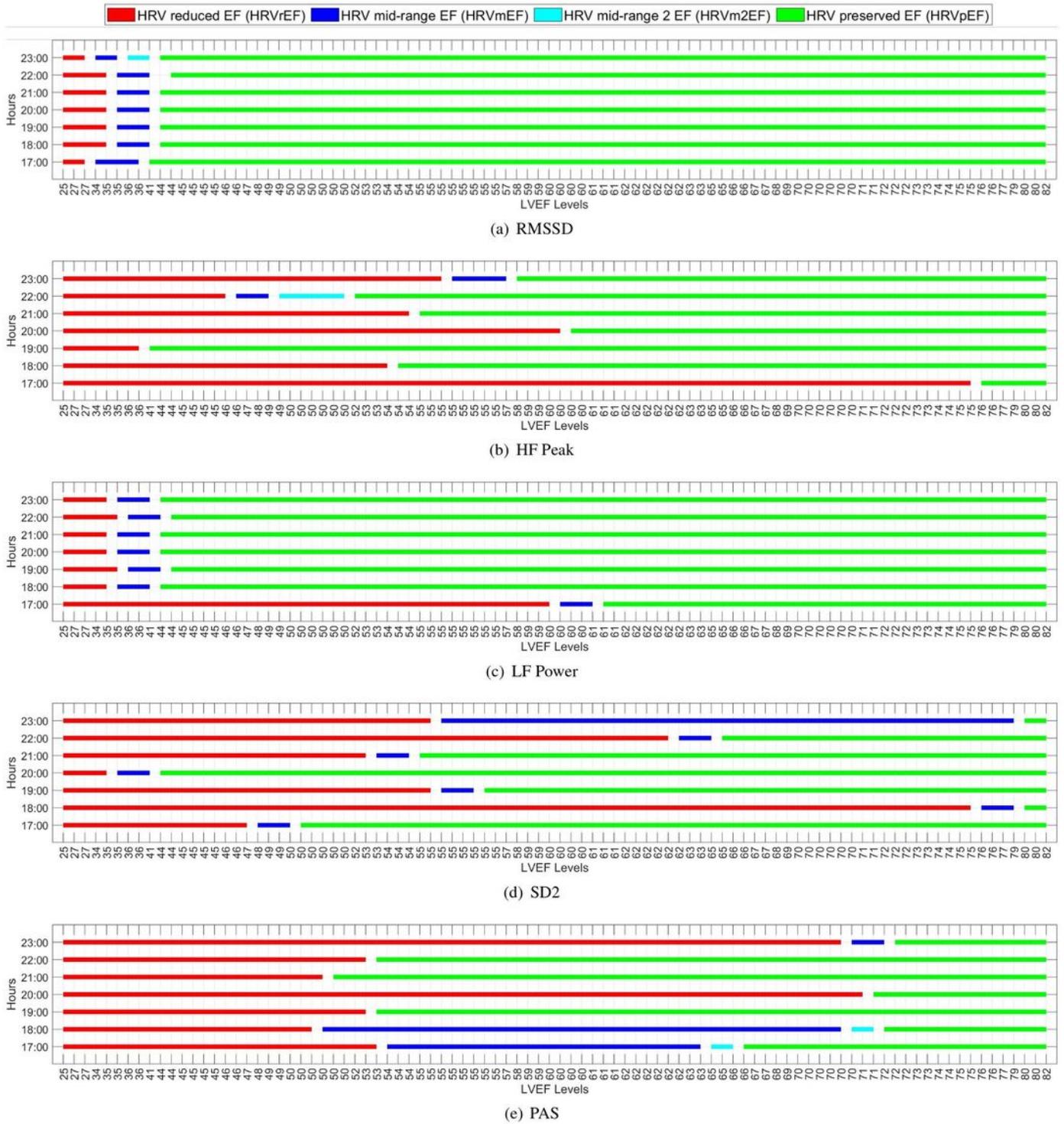


Figure 5

The evening (17:00-23:00) hours distribution of heart rate variability ejection fraction (HRVEF) groups across patients based on Jenks natural breaks.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFigureS1.pdf](#)
- [SupplementaryFigureS2.pdf](#)
- [SupplementaryTableS1.pdf](#)