THE EUROPEAN SOCIETY OF MEDICINE
Medical Research Archives, Volume 14 Issue 1

RESEARCH ARTICLE

An AI-Assisted Research Study Demonstrating the Importance of Shannon Entropy in Detecting Atrial Fibrillation through Heart Rate Time Series

Orazio A. BARRA[a,b]

[a]University of Calabria, Envir. Eng.Dept., Arcavacata, Cosenza, 87036, Italy [hereinafter UNICAL]
[b]International Polytecnic "Scientia et Ars", Feroleto Antico, 88040, Italy [hereinafter POLISA]

orazioantonio.barra@gmail.com

OPEN ACCESS

PUBLISHED
31 January 2026

CITATION
BARRA, OA., 2026. An AI-Assisted Research Study Demonstrating the Importance of Shannon Entropy in Detecting Atrial Fibrillation through Heart Rate Time Series. Medical Research Archives, [online] 14(1). https://doi.org/10.18103/mra.v14i1.XXXX

DOI
https://doi.org/10.18103/mra.v14i1.XXXX

ISSN
2375-1924

ABSTRACT

Prompt detection of Atrial Fibrillation is crucial to avert the serious complications linked to this arrhythmia. The diagnosis obtained from ECG-Holter Monitoring is unreliable unless the arrhythmia occurs during the course of these examinations. The paper presents a novel and robust methodology for the prediction and diagnosis of Atrial Fibrillation, employing Heart Rate Variability analysis of a patient, grounded in the most advanced techniques of the statistical mechanics of complex disordered systems, and suitable for integration into clinical practice. The methodology has also employed Artificial Intelligence (following an adequate period of Machine Learning) to verify the results via a secondary, independent process. The research is an observational study involving several thousand individuals who underwent experimental heart rate monitoring and subsequent variability analyses. Among the numerous markers evaluated in this analysis, four of them demonstrate the ability to detect and diagnose fibrillation with high sensitivity but limited specificity, and only if AFIB occurs during the monitoring period. Notably, one indicator, Shannon Entropy, exhibits exceptional performance by effectively detecting Atrial Fibrillation with both high sensitivity and specificity, and even if episodes occurred in the recent or distant past, demonstrating a significant "memory effect". This fact provides clinicians with an innovative approach for detecting and or predicting this important arrhythmia, even in the absence of ECG analysis, by solely monitoring the patient's heart rate over a 24-hour period. This approach substantially enhances the detection of AFIB episodes and facilitates the development of preventive measures and prophylactic therapies to mitigate the adverse effects of the arrhythmia.

1. Introduction

The efficacy of Heart Rate Variability (HRV) analysis for the early detection of cardiac pathologies has been demonstrated generally[1], along with the significance of the maximal entropy principle in the temporal correlations of HRV[2]. This study focuses on a specific cardiac condition, atrial fibrillation (AFIB), with the objective of identifying the most appropriate Heart Rate Variability (HRV) parameters capable of accurately diagnosing or predicting AFIB for direct implementation in clinical settings. and to surpass the limitations of current AFIB detection methods that predominantly depend on the patient's ECG readings.

Due to its propensity to trigger cardiovascular events[3-7] and subsequent cognitive decline[8], AFIB is a prevalent cardiac arrhythmia with potentially significant health consequences if not identified early and managed with preventive measures to reduce the associated risks. The potential development of blood clotting during the irregular rhythm of AFIB may result in subsequent cerebral strokes, or similar events, following the conclusion of the AFIB episode.

According to the most recent ESC Guidelines[9], AFIB is associated with a fivefold increased risk of ischemic stroke. Preventing stroke through oral anticoagulation is a critical aspect of managing patients with atrial fibrillation. The administration of anticoagulants - particularly the contemporary class of New Oral Anticoagulants (NOA) - significantly diminishes this risk. However, it is unfeasible to envisage the treatment of entire populations with lifelong anticoagulant therapy, due to the high costs associated with NOAs and their potential side effects, including bleeding events and alterations in vital clinical parameters.

Therefore, it is essential to differentiate individuals who are legitimately at risk of experiencing sporadic or recurrent AFIB episodes from those who are not. Unfortunately, the principal diagnostic techniques for AFIB, namely the ECG (electrocardiogram) and the 24-hour Holter Monitoring of cardiac activity, possess a limited probability of detecting the arrhythmia, as it must occur precisely during the testing period to be identified.

An alternative methodology capable of assessing the risk of AFIB in an individual should therefore be viewed as advantageous, both for increasing the detection of AFIB cases and for differentiating patients who most require anticoagulant therapy from others.

The primary objectives of this paper are therefore:

- To demonstrate that HRV analysis can provide a novel methodology for the early detection of AFIB by utilizing several indicators commonly employed in this form of analysis.[10]
- To consider the HRV-based methodology outlined above as a complementary or integrative approach to the traditional ECG-based method, thereby enhancing the identification of AFIB episodes.
- To present clinicians with a simple and effective tool - the Shannon Entropy - whose values, obtained from straightforward 24-hour heart rate recordings acquired through wearable cardio-frequency monitors, can rapidly distinguish individuals with, or at risk of, AFB from others.
- To enhance the effectiveness of risk stratification algorithms for predicting stroke and thromboembolism in patients with atrial fibrillation, such as the widely used CHA(2)DS(2)-VASc score[11] or the Framingham Risk Score[12] (commonly used for severe coronary heart disease) by incorporating the Shannon Entropy (ShE) score into the algorithm.

2. Methods

A total of about 7500 patients were included in the study, yielding a net effective count of 7315 valid patients after excluding records impacted by artifacts or errors in documentation. The patients have been divided into the following two groups and five subgroups:

- A-Group, comprising 6511 individuals with no prior or current episodes of AFIB, divided into two subgroups:

✓ A1 (1813 subjects) comprising healthy individuals (control group), defined by the absence of substantial disease and the lack of long-term therapy;

✓ A2 (4698 subjects A2 (4698 subjects) comprising individuals with diverse comorbidities, excluding AFIB.

- B-Group, comprising 804 patients with AFIB , divided into three subgroup as followss:

✓ B1 (247 patients): individuals with recent episodes of AFIB (within days or weeks prior to the HRV test, as ascertained from the patients' medical history;

✓ B2 (227 patients): individuals with a history of atrial fibrillation episodes occurred months prior to the HRV test, as ascertained from their medical records;

✓B3 (330 patients):individuals with atrial fibrillation episodes identified during the HRV test and confirmed by the recorded ECG trace.

The entire methodology employed in the development of HRV analysis is outlined in the following four steps:

STEP 1: FROM THE PATIENTS TO THE HEART RATE TIME SERIES

All patients have undertaken a 24-hour Heart Rate Monitoring (HRM), conducted using Mortara 12-channel Holter devices and/or high-quality wearable cardio-frequency monitors (such as POLAR V800, E-MOTION Faros, MINICARDIO HOSAND PROFESSIONAL, and similar equipment). From the HRM records, the heart rate time series , or RAW RR files, have been extracted - where RR represents the time interval between two consecutive heartbeats - and subsequently decoded, cleaned (to remove artifacts and malfunctions), and processed to produce RR files consisting of approximately 100,000 RR intervals within 24 hours for each subject, expressed in milliseconds, and prepared for subsequent HRV analysis.

STEP 2: ASSESSING THE PATIENT'S HEART RATE VARIABILITY USING A 50-MARKER VECTOR

Following the preparation of the RR interval files, MATLAB-based software has been developed and employed to perform a comprehensive HRV analysis for each patient. This analysis is based on the evaluation of a series of 50 distinct HRV indicators, providing a thorough characterization of the patient's HRV directly from the recorded RR interval data. TABLE 1 lists these indicators, or markers, together with their respective definitions. From the table, it is evident that the HRV of a subject can be characterized by three distinct categories of markers: linear markers in the time domain, linear markers in the frequency domain (obtained using Fast Fourier Transforms (FFT) and Autoregressive Methods, and assessed independently for Very Low, Low, and High Frequency Bands), and non-linear markers.

Table 1: List of hrv markers evaluated from rr "raw" data files

TYPE		MARKER (VARIABLE)	UNITS	DEFINITIONS
LINEAR VARIABLES	TIME DOMAIN	RR or NN interval	[ms]	*Time Interval between two consecutive QRS complexes*
		Mean RR	[ms]	*The Mean of RR Intervals over a 24 hours period*
		SDNN	[ms]	*Standard Deviation of RR Intervals (24h)*
		Mean HR	[b/m]	*The Mean Heart Rate (24h)*
		STD-HR	[b/m]	*Standard Deviation of Instantaneous Heart Rate Values (24h)*
		RMS-SD	[ms]	*Square Root of Mean Square Differences between Successive RR Intervals (24h)*
		NN50	[n]	*Number of Successive RR Interval Pairs that Differ > 50 ms (24h)*
		pNN50	[%]	*NN50 divieded by the Total Number of RR Intervals (24h)*
		HRV TIN	-	*Integral of the 24h RR Interval Histogram divided by its Height*
		Baseline TIN	[ms]	*Baseline Width of the RR Interval Histogram (24h)*
		Mean RR5	[ms]	*The Mean of the Average NN Intervals over 5 min periods*
		SDANN	[ms]	*Standard Deviation of the Average NN Intervals over 5 min periods*
	FREQUENCY DOMAIN*	Peak Frequency	[Hz]	*VLF, LF and HF Band Peak Frequencies, evaluated by FFT and AR Methods*
		Absolute Power	[ms^2]	*Absolute Powers of VLF, LF and HF Bands (both FFT and AR)*
		Relative Power	[%]	*Relative Powers of VLF, LF and HF Bands (both FFT and AR)*
		Normalized Power	[%]	*Powers of LF and HF Bands in Normalized Units [i.e.excluding VLF Band]*
		Total Power	[ms^2]	*Total Power*
		LF/HF	-	*Ratio Between LF and HF Band Powers*
NON LINEAR VARIABLES	Poincarè	SD1	[ms]	*Standard Deviation of Poincarè Plot [Rn+1 vs Rn] Orthogonal to the Identity Line*
		SD2	[ms]	*Standard Deviation of Poincarè Plot [Rn+1 vs Rn] Along the Identity Line*
	Recurrence Plot Analysis	RPL min	[b]	*Mean Line Length*
		RPL max	[b]	*Maximum Line Length*
		REC	[%]	*Recurrence Rate*
		DET	[%]	*Determinism*
		ShE	_	*Shannon Entropy*
	Others	ApEn	_	*Approximate Entropy*
		SampEn	_	*Sample Entropy*
		DFA α1	_	*Detrended Fluctuations Analysis: Short Term Fluctuation Slope*
		DFA α2	_	*DFA: Long Term Fluctuation Slope*
		D2	_	*Correlation Dimension*
		BMP	[%]	*"LIFE Potential" or "BARRA-MORETTI Potential"* [4]

** Evaluated Both through Fast Fourier Transform (FFT) and AutoRegressive Methods (AR), and, separately, for each band [Very Low Frequencies-VLF; Low Frequencies - LF; High Frequencies - HF]*

The latter category encompasses the Shannon Entropy (ShE)[13–18] and several other indicators of clinical significance. The three frequency bands utilized in the aforementioned FFT are: HF: 0.16–0.40 Hz; LF: 0.05–0.15 Hz; and VLF: 0.01–0.04 Hz.

The vector composed of these 50 marker values functions as a unique identifier for each individual, analogous to a sequence of their DNA. *However, unlike DNA, its components may fluctuate over time depending on the individual's age, health condition, and general well-being.*

STEP 3: GENERATING THE SEQUENCE OF MARKER ALARMS

By comparing the values of the vector components above to the average values of the same markers in the control subgroup A1 (healthy patients), it is possible to identify alarms for each marker, which occur whenever a marker's value significantly exceeds or falls below the corresponding average value of the control group. Moderate or severe alarms are triggered for each marker when the patient's marker value exceeds or falls below a factor of σ or 2σ, respectively (\σ representing the standard deviation of the distribution of the marker's values around its mean in the control group). Indicating in green the absence of alarms, in yellow the presence of a moderate alarm, and in red the presence of a severe alarm, each vector introduced in Step 2 becomes a sequence of markers' alarms, distinguished by an alternating pattern of green, yellow, and red colors. Within this framework, HRV can function as a reliable and autonomous diagnostic tool, assuming that the distinctive pattern of a pathology can be identified through the specific distribution of a patient's sequence of marker alarms.

STEP 4: PATHOLOGY IDENTIFICATION FROM THE PATIENT'S MARKER ALARMS

Based on the sequence of the patient's marker alarms, it becomes possible to identify the presence of a pathology once the characteristic alarm pattern associated with the pathology is recognized. *This paper's primary objective is to identify the "AFIB pathology" by establishing its characteristic pattern of marker values, which appear to be altered (alarms) relative to the normal values observed in the control group.*

Upon completing the procedures detailed in Steps 1, 2, and 3, several analyses have been conducted to identify the characteristic marker patterns and alarms for each of the previously mentioned groups and subgroups, aiming to uncover any significant clinical findings. The statistical significance of these results has been assessed by means of standard t-statistics and p-values where appropriate, as well as of relevant sensitivity and specificity percentages[19–36]. These analyses also incorporated an original graphical evaluation of the distribution of marker values within each group, as well as their overlap or distinctions between groups, to provide an easily interpretable and practical tool for clinicians. This graphic will be employed in the following section to illustrate the primary research findings.

STEP 5: VALIDATION OF THE PATHOLOGY IDENTIFICATION BY MEANS OF AI-ASSISTED PROCEDURE

Concurrently with the aforementioned algorithms utilizing conventional methods from the statistical mechanics of complex disordered systems, an Artificial Intelligence (AI) methodology has been developed to validate and substantiate the previously acquired results.

The RR-files have been directly utilized as inputs for a multi-layer feed-forward neural network, which has been pre-trained to facilitate patient classification. However, before training the network, preliminary procedures were carried out to confirm the effective number of markers via principal component analysis and to implement data augmentation owing to the wide scope of the input data. Having developed this basis, a machine learning (ML) approach has been employed to train the network and illustrate its proficiency in accurately classifying patients, distinguishing healthy individuals from those with atrial fibrillation.

The concordance between AFIB detection using the algorithms introduced in stages 1, 2, 3, and 4 and AFIB detection through AI-based methods exceeds 99%.This affirms the growing importance of AI in healthcare: in recent years, numerous devices and sophisticated algorithms have been effectively employed to assist medical professionals. One of the main goals of this collaboration between humans and machines is to guarantee widespread accessibility - particularly in low-income and remote areas - to medical services, as well as to reduce the time needed to achieve a diagnosis. Certainly, for AI-driven devices to

efficiently analyze large volumes of data and make timely precise decisions, they must first undergo proper training. *The exhaustive machine learning methodologies utilized in this study are thoroughly described in the aforementioned references 1 and 2.*

The aforementioned database, which consists of 7315 RR 24h-HRM patient records, comprises: 2214 records provided by the Ascoli Piceno Public Hospital between 2016 and 2020 as part of a cooperation agreement signed with POLISA and within a contract obtained from UNICAL; 1887 records downloaded by the Physiobank MIT online database[37] and 3214 patient records generated by POLISA's MATCH HRV NETWORK inside the Pythagoras Project, which is co-funded by REGIONE CALABRIA [see the "Acknowledgements" below].

The amounts and types of files received by each source are summarized in **Table 2**, which is divided into three categories: "HP" (Healthy People), which refers to patients with no pathologies and no ongoing therapies; "AFIB," which refers to patients with atrial fibrillation; and "ANY PATHOLOGY BUT AFIB," which refers to patients with any pathology other than AFIB.

Table 2: Sources of the analyzed rr "raw" data files

	HP	AFIB	All Pathologies But AFIB	TOTAL
FROM THE ASCOLI PICENO HOSPITAL	756	700	758	2214
FROM THE PHYSIOBANK MIT ON-LINE DATABASE	57	104	1726	1887
FROM THE MATCH HRV-NETWORK Managed by POLISA	1000	0	2214	3214
GRAND-TOTAL:	1813	804	4698	7315

3. Results

The findings indicate that, although four HRV markers demonstrate high sensitivity in detecting AFIB episodes - albeit exclusively those occurring during the HRM performed for HRV analysis - they exhibit limited specificity in diagnosing AFIB. Consequently, these four markers will be designated as "*second-choice AFIB markers*." However, a fifth marker, Shannon Entropy (ShE), appears to provide both high sensitivity and specificity in the diagnosis of AFIB. Furthermore, the ShE demonstrates a peculiar and atypical behavior: it appears capable not only of detecting an active AFIB episode occurring during the HRM of the HRV test but also of retaining evidence of AFIB episodes that took place weeks or months earlier. This implicitly encompasses the associated risk of recurrent AFIB episodes in the near to medium-term future, which correlates proportionally with the number of prior AFIB episodes. Therefore, the ShE will be designated as the "*First Choice AFIB Marker*." Within this framework, the principal research findings can be summarized into two key points: the first concerning "**the 2nd choice AFIB markers" (Sub-section 3.1**), and the second relating to "**the 1st choice marker" (Sub-section 3.2)**. Additional statistical observations will complete the results section (**Subsection 3.3**).

3.1 FIRST KEY-POINT: "THE 2nd CHOICE AFIB MARKERS"

These markers include the non-linear variable SD1 as described in TABLE 1, along with the three linear variables pNN50, Mean RR5, and the ratio (RMS-SD)/(MEAN RR). All these Markers increase their values during AFIB episodes. Their behavior is shown in details, numerically in **TABLE 3**, upper section, and graphically in **FIGURE 1.**

All marker values are normalized to the average values observed in healthy individuals to achieve dimensionless units; thus, the values in the table denote the ratio of the absolute marker values to the corresponding average values for healthy individuals (HP), specifically those in the A1-Sub-Group [Control Group]. In this normalization, the mean value of the marker for the HP is uniformly established at "1," with the range of HP values characterized by the Standard Deviation (SD) relative to unity, often ranging from a few percent to several tenths of a percent.

Table 3. Typical range of values and their principal statistical characteristics for the markers addressed in the 1st and 2nd key points

		A1 SUBGROUP			A2 SUBGROUP			B3 SUBGROUP		
UPPER SECTION	HRV MARKER	Range of values	Average Value	Standard Deviation	Range of values	Average Value	Standard Deviation	Range of values	Average Value	Standard Deviation
	RMS-SD / Mean RR	0.38-1.71	1,00	0,33	0.38-3.14	1.76	0.69	1.23-2.86	2.05	0.41
	pNN50	0.26-1.30	1,00	0.26	0.26-2.14	1.20	0.47	1.53-4.05	2.79	0.63
	SD1	0.36-1.62	1,00	0.31	0.36-3.87	0.99	0.88	0.96-2.51	1.73	0.39
	VN(MEAN)RR5	0.43-1.28	1,00	0.21	0.21-2.19	1.20	0.50	1.23-2.51	1.87	0.32
	ShE	0.93-1.17	1,00	0.06	0.93->2	1.5	0.3	see below		
LOWER SECTION	HRV MARKER	B1 SUBGROUP			B2 SUBGROUP			B3 SUBGROUP		
		Range of values	Average Value	Standard Deviation	Range of values	Average Value	Standard Deviation	Range of values	Average Value	Standard Deviation
	Shannon Entropy	0.79-0.93	0.86	0.03	0.71-0.87	0.79	0.04	0.02-0.78	0.40	0.19
NOTE:	Alla values are dimensionless as they are divided by the corresponding average value of the A1 (Healthy) Group.									

In Table A , the table's parts are indicated in the first column, and the marker names - including the AFIB second-choice markers taken into consideration in this subsection as well as the first-choice marker shown in the following subsection 3.2 - are listed in the second column. Progressing from left to right, three sets of columns are displayed, corresponding to the A1, A2, and B3 subgroups of patients in the upper section of the table, and to the B1, B2, and B3 subgroups in the lower one. Each set of columns is divided into three columns, each representing the range of values, the mean value, and the standard deviation (around the mean) of the markers for each subgroup, respectively.

The measured results for the four 2nd choice markers are subsequently presented in a more clear and practical graphical format in FIGURE 1. The four markers mentioned above are depicted within adjacent, distinct vertical blocks. Within each block, the vertical bars depict the complete distribution of marker values among patients in the A1, A2, and B3 subgroups. In contrast, the B1 and B2 subgroups are not depicted, as they lack members because they exhibited no significant alterations in the values of the four markers for patients within these categories.

Although these four markers seem adequately capable of identifying the presence of AFIB, two main obstacles limit the implementation of these markers in clinical practice:

- The markers are capable only of detecting ongoing AFIB episodes, that is, episodes occurring during the HRM (precisely as occurs with ECG and conventional Holter Monitoring), and do not appear to possess the ability to assess previous AFIB episodes (the patients in the B1 and B2 categories do not show changes for these markers, as mentioned earlier)

Due to the apparent overlap in the range of values presented in the second and third sets of columns in Table 3 [or, visually, in the corresponding A1, A2, and B3 vertical bars in Figure 1], most of the values for these four markers are shared across numerous other pathologies (A2) and, to some extent, with the Healthy People Subgroup (A1). For instance, an SD1 Marker value of 1.23 is detected in the A1 and A2 populations, as well as in the B3 population. Therefore, the specificity of these markers for AFIB detection seems to be relatively limited, except for the marker values within the uppermost zone of the B3 vertical bars in Figure 1, which will be more precisely detailed in the subsequent subsection 3.3.

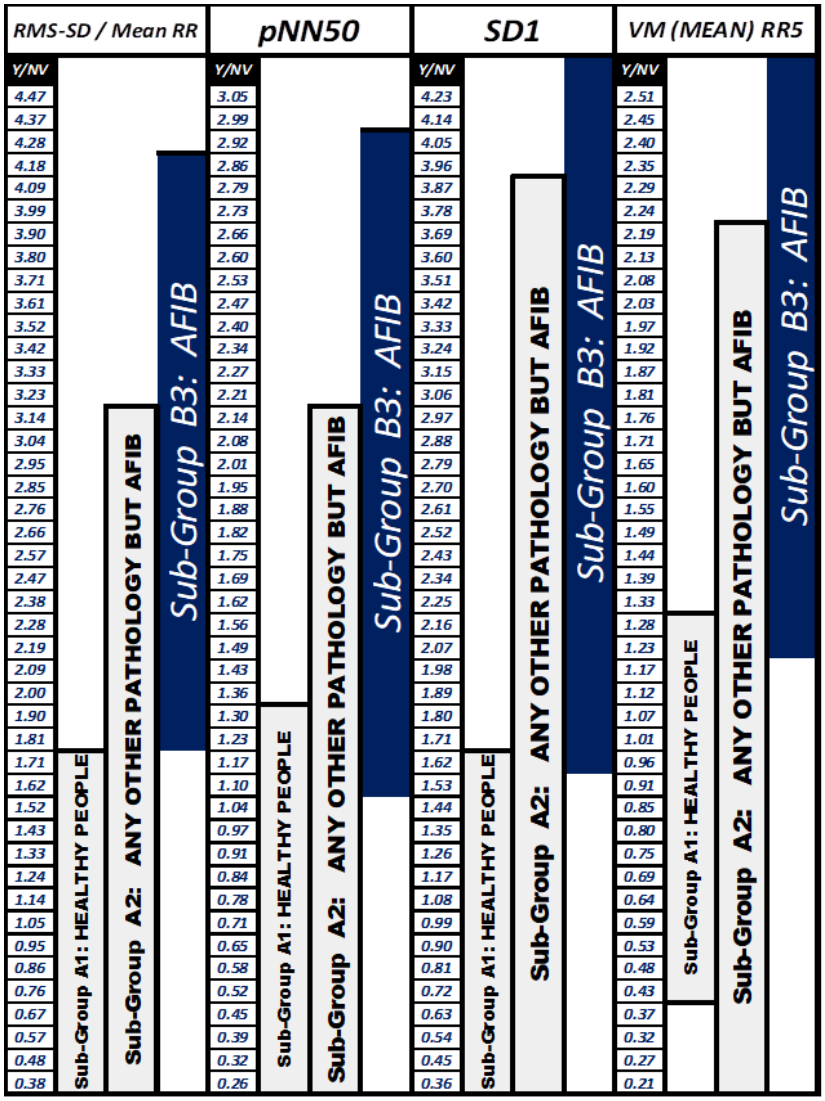

Figure 1: Spreading of the 2nd choice markers values for different sub-groups of patients

3.2 SECOND ITEM: THE 1ST CHOICE AFIB MARKER: THE SHANNON ENTROPY (SHE)

The impressive performance of the ShE in AFIB detection is demonstrated in the preceding **TABLE 3, lower section**, in numerical format, and in **FIGURE 2,** in graphical format, utilizing the same presentation style as that used for the markers in Figure 1. The units on the y-axis are once again normalized relative to the mean ShE value observed in healthy individuals. Specifically, the y-axis value indicates the ratio of the ShE value obtained for the subject under investigation to the average ShE value measured across subjects within the A1 cohort. The sole significant distinction between Figures 1 and 2 is that all subgroups B1, B2, and B3 are now comprehensively integrated and adequately represented in Figure 2.

The results depicted in the figure possess substantial clinical significance for the diagnosis of AFIB, as they categorize the entire patient population into two distinct, well-defined groups: individuals with a ShE normalized value below 0.93 and those with a value equal to or exceeding 0.93. This signifies that:

❖If a patient demonstrates a normalized ShE value of 0.93 or higher (graphically positioned above the horizontal dashed line), indicating a level exceeding 93% of the minimum ShE value classified as "normal", the patient is neither suffering from AFIB nor at considerable risk of future AFIB episodes. Therefore, the patient is currently ineligible for preventative measures targeting AFIB-related problems. More specifically, if the ShE value lies within the range of 0.93 to 1.16, the patient is regarded as healthy or potentially impacted by conditions not associated with AFIB. If the value surpasses 1.16, the patient is considered unhealthy and is influenced by underlying conditions other than AFIB. including likely carcinomas, congestive cardiac failures, severe neurological disorders, and others, but the clinical significance of elevated ShE values is beyond the scope of this study.

❖If a patient's normalized ShE value falls below 0.93 (below the horizontal dashed line), indicating

less than 93% of the minimal "normal" ShE value, the patient is highly likely to experience episodes of atrial fibrillation (AFIB). Consequently, the patient is eligible for prophylactic treatments (e.g., non-vitamin K antagonist oral anticoagulants, or NOACs) to reduce the risk of AFIB-associated complications. More specifically:

- If the ShE value is within the range of 0.86-0.93, it indicates that the patient most likely experienced symptomatic or asymptomatic atrial fibrillation episodes in the distant past (one or more months prior).
- If the value falls within the range of 0.78 to 0.86, it suggests that the patient likely experienced AFIB episodes in the recent past (days or weeks) and/or in the distant past.
- If the value decreases between 0.70 and 0.78, the patient experienced AFIB episodes either in the recent past or during the 24-hour HRM recording during the HRV analysis.

If the value is below 0.78, the patient demonstrated episodes of AFIB during the 24-hour HRM recording conducted for the HRV test.

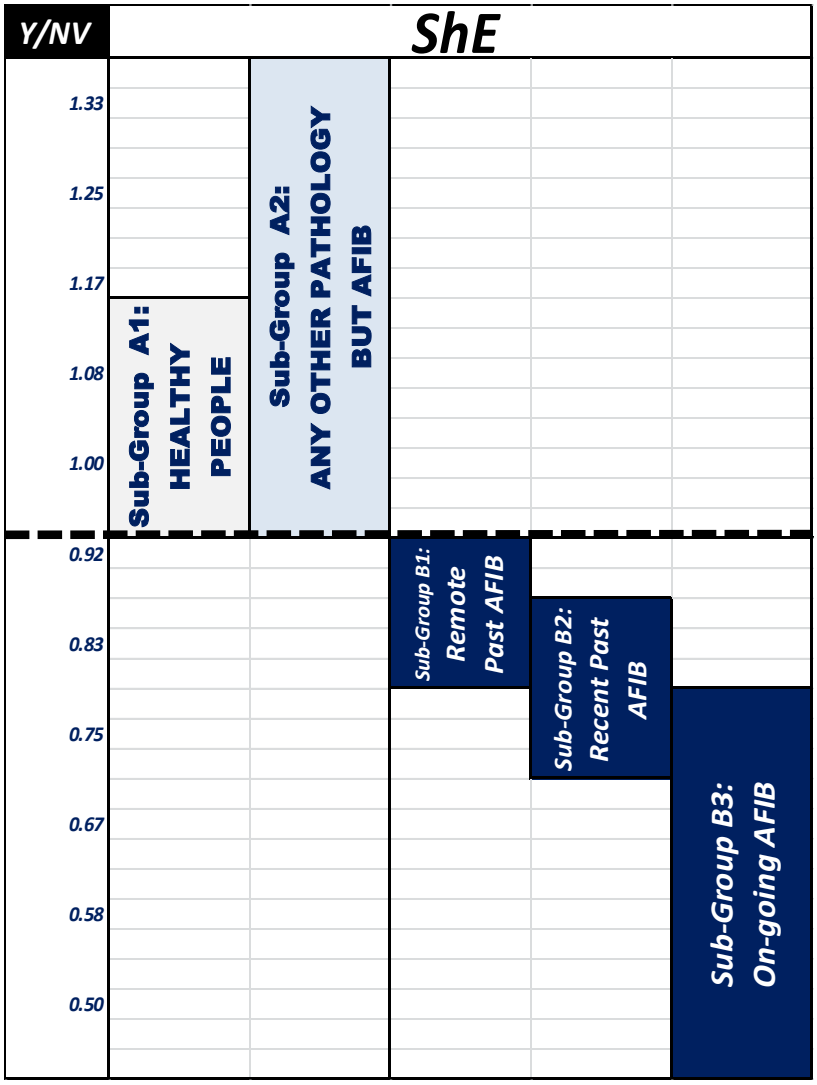

Figure 2: Spreading of the she measured values for the five sub-groups of patients

As depicted in Figure 2, the two main limitations of the prior markers presented in sub-section 3.1 have been resolved, and the use of ShE in clinical practice for AFIB prediction and detection is now prepared for deployment; indeed:

- The ShE is capable not only of detecting AFIB episodes occurring during the HRM but also of demonstrating a notable "memory effect" concerning AFIB episodes that took place in the recent (B2-Subgroup) or distant (B1-Subgroup) past. When a patient exhibits a ShE value within the range corresponding to one of the B Subgroups, even if no AFIB episodes are documented during the HRM, it is crucial to evaluate the likelihood of future AFIB occurrences and to consider initiating preventive treatments to reduce AFIB-associated risks.

• Despite the limited specificity of the secondary AFIB markers in Subsection 3.1, the ShE demonstrates both high sensitivity and specificity. This assertion is visually corroborated by Figure 2, which clearly shows that the region occupied by the ShE values in AFIB conditions is distinctly and entirely separated from the areas associated with the ShE values of healthy individuals and/or those with other pathologies, as further quantified in the subsequent subsection 3.3.

For the reasons outlined above, the ShE should be considered the primary diagnostic tool for AFIB, offering clinicians a valuable means to distinguish individuals at risk of AFIB from others. Once the ShE drops below 93% of the average ShE value observed in healthy individuals, atrial fibrillation can be accurately diagnosed or anticipated. In other words, low ShE values appear to be a prerequisite for the occurrence of AFIB episodes, and the lower the ShE value, the greater the probability of AFIB episodes.

3.3 ADDITIONAL STATISTICAL OBSERVATIONS

A. **P-VALUE**: The single p-value of the entire B group of patients compared to the entire A group has been taken into consideration because the current paper is solely concerned with the detection of the AFIB. The p-value for the 2nd choice markers was $p<0.05$ while for the ShE the p-value was <0.001, or nearly zero. Because Groups A and B were clearly split in Figure 2, this was also visually understandable. and consistent with the amount of degrees of freedom (about 800) because of the large number of patients examined, which is uncommon in these kinds of investigations.

B. **SENSITIVITY AND SPECIFICITY**: **TABLE 4** presents the sensitivity and specificity values for both the primary and secondary markers previously examined. *For the second choice markers,* TABLE 4 indicates that their sensitivity in AFIB detection is satisfactory. (σ > 95%), resulting in a very low proportion of "*false negatives FN*" " (FN = 1 – σ < 5%). However, due to the significant overlap in the range of values already observed in the second and third columns of Table 3 and in Figure 1, most of the high values of these four markers are also present in numerous other pathologies and consequently, the specificity of these markers for AFIB detection appears to be quite limited., with s values in the range 25-33%, and therefore an high percentage of "False Positive" [FP = 1-s], which is the range 67-75%.. The four markers show an appropriate specificity in AFIB identification only within the narrow range of marker values that surpass the top limit of the range of values indicated by the marker for the Subgroup A2 (*Table 4, right section*)., with s values in the range 65%-70% and a percentage of "False Positive" [FP = 1-s], in the range 30%-35%.

For those the first choice marker is concerned, when applied in clinical practice, ShE appears to demonstrate both high diagnostic sensitivity (σ >90%), resulting in a very low percentage of false negatives (FN = 1 - σ < 1%), and very high diagnostic specificity (s>99.9%) leading to a very low percentage of false positives (FP = 1 - s ≤ 0.1%). This assertion is supported by Figure 2, which clearly demonstrates that the region occupied by the ShE values in AFIB conditions is distinctly separated from the areas associated with the ShE values of healthy individuals and/or those affected by other pathologies: once the ShE value falls below 93% of the average ShE value observed in healthy individuals, AFIB can be reliably diagnosed or predicted.

TABLE 4: Sensitivity σ and specificity s of the 1st and 2nd choice markers for AFIB detection

MARKER TYPPE	HRV INDEX	SENSITIVITY "σ"		SPECIFICITY "s"							
		σ	FALSE NEGATIVE	MARKER VALUE	s	FALSE POSITIVE	MARKER VALUE	s	FALSE POSITIVE
1st CHOICE MARKERS	ShE	> 99%	< 1%	<0,93	> 99,9%	< 0,1%	≥0.93	> 99,9%	< 0,1%
2nd CHOICE MARKERS	RMS-SD / Mean RR	> 95%	< 5%	≤3.32	< 30%	> 70%	>3.22	> 65%	< 35%
	pNN50	> 95%	< 5%	≤2.20	< 33%	> 67%	>2.20	> 69%	< 31%
	SD1	> 95%	< 5%	≤3.95	< 28%	> 72%	>3.95	> 67%	< 33%
	VN(MEAN)RR5	> 95%	< 5%	≤2.23	< 25%	> 75%	>2.23	> 70%	< 30%

Some concluding remarks based on the thorough analysis of the RR data files are as follows:

➢ No significant differences in the ShE values have been observed between males and females.

➢ Short-duration AFIB episodes lasting less than 6 minutes did not modify the She values: this finding seems to endorse the hypothesis that brief episodes of AFIB are inherently less hazardous than prolonged episodes in relation to the likelihood of precipitating a cardiovascular event.

➢ The alterations in SHE values are inadequate to differentiate between Atrial Fibrillation and Flutter, which, from this standpoint, may be regarded as the same clinical entity.

4 Discussion

ShE is the sole marker exhibiting substantially below-normal levels in cases of AFIB, whereas all other analyzed 49 markers remain unchanged or exhibit rapid increases. The same SHE values, when associated with other pathologies (including major conditions such as carcinomas, congestive heart failures, genetic neuropathies, etc.), tend to elevate and do not diminish. This unexpected finding adds to the ongoing important discourse concerning the connection between physics and information: the operational equivalence of information gain and entropy reduction has been widely acknowledged. Nonetheless, the notion that a subjective measure such as information could influence the "objective" thermodynamic properties of the system continues to be a matter of considerable debate. Nevertheless, it is difficult to deny that the process of information acquisition can be directly associated with the ability to perform effective tasks. Consequently, questions concerning thermodynamics, the second law, and the arrow of time have become intertwined with a longstanding dilemma: the problem of measurements in quantum physics. The specter of information casts a shadow over the sciences. Thermodynamics, the foundation of statistical mechanics, the quantum theory of measurement, the physics of computation, and numerous subjects within the disciplines of dynamical systems, molecular biology, genetics, and computer science all revolve around the shared concept of information. Well-defined indications exist concerning the significance of information within the domain of physics, as well as its function as a bridge between the natural sciences and the science of computation. Furthermore, for individuals aiming to attain proficiency in Machine Learning (ML) within the artificial Intelligence (AI) domain, understanding Shannon's Entropy is crucial, as ShE forms the basis of functions that comprise the essential tools for an ML practitioner [1,2]. Participating in this intricate debate surpasses the scope of this article; consequently, only a few concise considerations will be offered solely to promote new interdisciplinary perspectives between physics and medicine.

The ShE (X) is the total amount of information in an entire probability distribution of a set of n events Xi, defined as:

$$ShE(X) = -\Sigma_i P(x_i).\ \log_b P(x_i) = \Sigma_i P(x_i).\ \log_b [1/P(x_i)] = \Sigma_i P(x_i).\ \log_b I(x_i) \quad [1]$$

where I(Xi) =1/p(Xi) and p(Xi) are respectively the information content of event Xi and its probability to happen. The unit is the **nat = (1/0,69215) bit and usually the log base b is 2**. Therefore, greater is the quantity of information (and then ShE), greater is the number of bits necessary to represent it, and smaller is the probability to be generated and the possibility to forecast the event, and vice-versa.

At the same time the Thermodynamic Entropy ThE (hereinafter simply ThE) is a state function whose variation Δ(ThE)is defined as:

$$\Delta(ThE) = \Delta Q/T \quad [2]$$

being ΔQ = the heat given (-) or taken (+) by a system to/from a source at a temperature T (K). Given a thermodynamic system composed of 2 sub-systems, H and C, in contact with each other, the first at a Temperature TH greater than the temperature TC of the second one, heat spontaneously flows from the hotter to the colder subsystem and the total variation Δ(ThE) of the system is $\Delta(ThE) = -\Delta Q/TH + \Delta Q/TC > 0$. Therefore in a spontaneous process the ThE always increases: this implies that nature evolves towards a continuous growth of its global ThE in the Universe. It can be shown that:

$$ThE = k.\log D = -k.\log(1/D) \quad [3]$$

being k the Boltzmann constant = $1{,}38.10^{-23}$[J/K], and D a quantitative measure of the atomic disorder of the system. Therefore, greater is the disorder (and then ThE), greater is the probability to be generated, and vice-versa. It is evident that both ThE and ShE capture increasing randomness and that:

$$ShE \approx -\beta.ThE \quad [4]$$

where **β** is a dimensional constant, necessary to compensate the different dimension of the two entropies. **Disorder increases when information decreases**, that is **ShE conceptually is ,** dimensions apart, **a quantity of ThE with the sign changed ,** a sort of **"negative thermodynamic entropy" .**

It follows that a living organism continuously increases its own ThE and should fastly reach its state of maximum ThE , which is its death. How can it avoid this dangerous decay? The obvious answer is: eating, drinking, breathing, thus through its "metabolism", which, from the greek word μεταβαλλειν , means: "to exchange". But "To Exchange What?" *MATTER* ? No (what could be the meaning of an exchange of an atom of sodium or oxygen with another one?!); *ENERGY* ? No (1 kcal = 1 kcal, and human body takes from food & environment the same energy lost towards the environment); *ENTROPY*? Yes, and then *INFORMATION*!

The organism can only survive by continuously extracting "Negative ThE" (i.e., ShE) from the environment; the objective is to equilibrate all unavoidable ThE production with a constant influx of new ShE. It extracts "Food" from the environment (high atomic-molecular order > high ShE > low ThE) and returns the "Final Wastes" of its metabolism to the environment (completely degraded matter, low atomic-molecular orders > high ThE > low ShE): the balance is satisfied because the total thermodynamic system "organism + environment" increases its ThE, and the life of the system is saved. In this logic, AFIB, a status which implies a reduction of ShE, seems to indicate "a shortage of information" generated by a dys-autonomia - a condition in which the autonomic nervous system(ANS) does not work properly – possibly to compensate a previous excess [as it could be the case of a paroxysmal AFIB episode after an overeating period], or to indicate the need to be compensated, by means of high information content intake such as drugs or electric defibrillation or ablation[38,39] [as it could be the case of AFIB on pathological basis: valvulopathies, congestive heart failures, etc.].

All the above confirms the value and the meaning of the Shannon Entropy also for the information content present in the HRV-Markers, where it seems to have a behavior very similar to the ThE and to the Exergy (Gibbs Free Energy) in Energetics. AFIB is the one with the minor possible information content, and therefore, the easiest to be generated. It seems that AFIB is the "most likely condition" - i.e. a condition which can be reached without "efforts", or, in any case, towards which the organism spontaneously evolves in absence of constraints and consumption of work. On the contrary, what is usually considered "the normal condition" – thus the sinus rhythm – appears as a "steady state" where the organism must be constrained through an expensive equilibrium between the vagal and the sympathetic nervous systems, but remaining always ready to collapse towards AFIB once the equilibrium above, for any reason, is lost.

Although the above discussion is a novel approach to AFIB, it may be able to change the "roots" of the entire strategy to address it. However, as the next section also reports, more focused experimental evidence and more robust theoretical models are required.

5 CONCLUSION

A new and effective mechanism for predicting and diagnosing atrial fibrillation has been introduced. This methodology is ready for prompt incorporation into clinical practice and aims to exceed the constraints of conventional ECG and Holter monitoring methods. The methodology employs a 24-hour recording of a subject's heart rate time series using high-quality wearable cardio-frequency meters or comparable devices. Subsequent examination of heart rate variability is conducted using 50 markers calculated by advanced algorithms based on the statistical mechanics of complex disordered systems, with results validated by artificial intelligence following a sufficient machine learning period. Four markers demonstrate a high sensitivity for identifying atrial fibrillation (AFIB), albeit their specificity is limited, and they are effective only if AFIB occurrences occur during the heart rate recording period. Conversely, a fifth marker, Shannon Entropy, Shannon Entropy demonstrates outstanding sensitivity and specificity in AFIB detection, as well as a notable ability to identify AFIB regardless of whether events occurred during the monitoring period or in the recent or distant past. This information will allow the clinician to determine the patient's likelihood of recurrent atrial fibrillation

(AFIB) episodes, thereby facilitating preventative measures and preemptive therapies to mitigate the negative impacts of both sporadic and persistent AFIB episodes.

Furthermore it must be emphasized that the proposed ShE – based methodology:

➢ Permits the assessment of the simpler RR intervals to be performed instead of the PQRST complexes of the ECGs.

➢ Can utilize advanced wearable technology, such as high quality cardio-frequency meters, pulse oximeters, health tracking smartwatches or specially designed devices, to continuously record heart rate throughout the year, 24 hours a day, ensuring accurate diagnosis of AFIB episodes whenever they occur.

➢ Shows no significant difference in the values of relevant HRV markers between males and females.

➢ Has demonstrated that AFIB episodes lasting less than 6 minutes do not modify the relevant ShE marker levels, supporting the assumption that short-term AFIB is significantly less hazardous than long-term AFIB in terms of the potential to induce a cardiovascular event.

➢ Does not differentiate between Atrial Fibrillation and Flutter, which, from this perspective, may be regarded as the same clinical condition.

It must be underlined that *Shannon Entropy is the sole HRV measure that exhibits a large and unmistakable drop during AFIB circumstances*, whilst all other markers either remain constant or increase. Conversely, *AFIB is the only pathology that can diminish ShE values from normal levels*, as all other pathologies can only elevate ShE values. This unique characteristic establishes a peculiar double link between AFIIB and ShE, which, through the ShE, creates a connection between AFIB and the Thermodynamic Entropy. This allows one to view AFIB from a new perspective, as a natural status where the human body precipitates to, due to the second principle of thermodynamics, unless the sinus rhythm is maintained by means of adequate intake of energy (from food, environment, and so forth) capable of compensating the growing thermodynamic entropy and consequently the declining ShE. This is a hypothesis that may be better investigated mathematically developed mathematical modeling in the near future. Regardless, Shannon Entropy is always the main HRV metric for AFIB detection, allowing for the identification of AFIB even while other marker values remain unchanged. However, in circumstances where the diagnosis is ambiguous, alarms in other relevant markers might still be helpful.

Ethics

The research pertains to observational studies that do not incorporate innovative pharmaceuticals or intrusive techniques. The collection and transfer of data will not disrupt individuals' regular activities. Data were processed in a completely anonymous manner and in strict compliance with the principles established in the "Declaration of Helsinki" (1964), with the informed permission of the study participants. Nevertheless, the Ascoli Piceno Hospital, which supplied the data records referenced in this paper, has sought formal approval for the study from the pertinent Ethical Committee. The Ethical Committee of the Regione Marche gave this sanction to the Hospital on June 9, 2016, before the study commenced.

Conflict of Interest Statement:

The author reports no conflict of interest concerning the materials or methods used in this study or the findings presented in this paper.

Funding Statement:

The study object of this paper is forming part of two Research Projects named "MATCH-PYTHAGORAS" and "PYTHAGORAS 2.0", co-financed by the REGIONE CALABRIA, within the POR-CALABRIA 2014-2020, CUP (Unified Project Code): J48C1700025000 ; our thanks our due to this Italian Regional Authority, because, without their support, these results could not have been achieved.

Acknowledgements:

Our thanks are due to the Ascoli Piceno Hospital, URCC (Unità di Ricerca Clinica Cardiologica) directed by Dr. L. Moretti, and Arrhythmology Unit coordinated by Dr. P.Marchese, for the always accurate collection and supply of clinical data and for the helpful discussions, and to Dr. L. Vento for his competent and significant support in data handling and management.

ORCID ID:

Orazio A. BARRA: 0000-0003-0823-4488

References:

1. Barra O.A. et al. : Detecting cardiac pathologies via machine learning on heart-rate variability time series and related markers; *Scientific Reports* **10**, 8845 (2020) ; DOI: https://doi.org/10.1038/s41598-020-64083-4

2. Barra O.A. et al.: Analysis of temporal correlation in heart rate variability through maximum entropy principle in a minimal pairwise glassy model"; *Scientific Reports* **10**, 15353 (2020); DOI: https://doi.org/10.1038/s41598-020-72183-4

3. Hart, R.G.&Halperin, J.L. Atrial fibrillation and stroke: concepts and controversies. *Stroke* **32** (3), 803-808 (2001)

4. Lip, G.Y.& Lim, H.S. Atrial fibrillation and stroke prevention. The Lancet Neurology 6 (11), 981-993 (2007)

5. Binici, Z., Intziakilis, T., Nielsen, O.W., Kober, L., Sajadieh, A. Excessive supraventricular ectopic activity and increased risk of atrial fibrillation and stroke. *Circulation* **121** (17), 1904 (2010)

6. Albers, G.W. Atrial fibrillation and stroke: three new studies, three remaining questions. *Archives of Internal Medicine* **154** (13), 1443-1448 (1994)

7. Ezekowitz, M.D. et al. Warfarin in the prevention of stroke associated with non-rheumatic atrial fibrillation. *New England Journal of Medicine* **327** (20), 1406-1412(1992)

8. Knecht, S.et al.: Atrial fibrillation in stroke-free patients is associated with memory impairment and hippocampal atrophy. *European Heart Journal* **29** (17), 2125-2132 (2008)

9. 2024 ESC Guidelines for the management of atrial fibrillation, Clinical Practice Guidelines (2024)

10. Task force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology: Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation* **93** (5), 1043–1065 (1996)

11. Olesen JB, Lip GY, Hansen ML, et al. Validation of risk stratification schemes for predicting stroke and thromboembolism in patients with atrial fibrillation: nationwide cohort study. *BMJ.* 3**42** (2011) , doi: 10.1136/bmj.d124. PubMed ID: 21282258

12. D'Agostino, R.B., Vasan R.S. et al. (2008). "General cardiovascular risk profile for use in primary care: the Framingham Heart Study"., *Circulation*, 117(6), 743-53 (2008)

13. Barra, O.A. & Moretti, L. The "Life Potential": a new complex algorithm to assess "Heart Rate Variability" from Holter records for cognitive and diagnostic aims. Preliminary experimental results showing its dependence on age, gender and health conditions. Preprint at https://arxiv.org/abs/1310.7230 (2013)

14. Shannon, C.E.& Weaver, W. The Mathematical Theory of Communication, Book *Edited by University of Illinois Press*, *Urbana* (1949, republished on 1998)

15. Jaynes, E.T. Information Theory and Statistical Mechanics. *Phys. Rev.* **106**, 620, DOI: https://doi.org/10.1103/PhysRev.106.620 (1957)

16. Zurek, W.H. Complexity entropy and the Physics of information. *The Proceedings Of The Workshop On Complexity, Entropy, And The Physics Of Information* Held May-June, 1989 In Santa Fe, New Mexico, Volume VIII , Santa Fe Institute Studies in the Sciences of Complexity, CRC Press (2018)

17. Cerf,N.J. &Adami, C. Negative Entropy and Information in Quantum Mechanics. *Phys. Rev. Lett.* **79**, 5194 (1997)

18. Gappmair, W. Claude E. Shannon: the 50th anniversary of information theory. *IEEE Communications Magazine* **37**, 4 , 102-105 DOI: https://doi.org/10.1109/35.755458(1999)

19. Goedeken, E.A. A Mind at Play: How Claude Shannon Invented the Information Age. *Information & Culture: A Journal of History, University of Texas Press* **53**, 3/4 , 377-378 (2018)

20. Steeb, W.H. A Handbook of Terms used in Chaos and Quantum Chaos, *BI-Wissenschaftsverlag, Mannheim*, (1991).

21. Zaslavsy, G.M. Chaos in Dynamic Systems, *Harwood, Chur* (1985)

22. Mandelbrot B.B., The Fractal Geometry of Nature, *Freeman, San Francisco* (1982)

23. Yamamoto, Y,. & Hughson, R.L. Coarse-graining spectral analysis: new method for studying heart rate variability. *J Appl Physiol.* **71**, 1143-1150 (1991)

24. Babloyantz, A &Destexhe, A. Is the normal heart a periodic oscillator? *BiolCybern.* **58,** 203-211 (1988)

25. Saul, JP, Albrecht, P, Berger, R.D., Cohen, R.J. Analysis of long-term heart rate variability:

methods, 1/f scaling and implications. In: *Computers in Cardiology 1987. IEEE Computer Society Press* .Washington, DC; 419-422 (1988)

26. Brennan, M,Palaniswami, M, Kamen, P.Do existing measures of Poincaré plot geometry reflect non-linear features of heart rate variability? *Proc. IEEE Transactions on Biomedical Engineering,* **48**, 1342-1347 (2001)

27. Kanters, J.K., Holstein-Rathlou, N.H., Agner, E. "Lack of evidence for low-dimensional chaos in heart rate variability". *Journal of Cardiovascular Electrophysiology* **5**, 7, 591–601 (1994)

28. Storella, R.J. et al."Approximate entropy and point correlation dimension of heart rate variability in healthy subjects". *Integrative Physiological & Behavioral Science* **33**, 4, 315–20 (1994)

29. Carrasco S., M.J. Cait´an, R. Gonz´alez, and O. Y´anez. Correlation among Poincar´e plot indexes and time and frequency domain measures of heart rate variability. *J MedEngTechnol*, **25**, 240–248 (2001)

30. Kobayashi ,M, Musha,T. 1/f fluctuation of heart beat period. *IEEE Trans Biomed Eng.* **29**, 456- 457 (1982)

31. Guzzetti, S. et al. Non-linear dynamics and chaotic indices in heart rate variability of normal subjects and heart-transplanted patients. *Cardiovascular Research* **31**, 441–446 (1996)

32. Lake, D.E., Richman, J.S., Griffin, M.P., Moorman, J.R., Sample entropy analysis of neonatal heart rate variability. *Am.Jour.of Phys.* **283**, R789–R797 (2002)

33. Penzel T., Kantelhardt J.W., Grote L., Peter J-H., and Bunde A., Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea. *IEEE Trans Biomed Eng.* **50**, 10, 1143–1151 (2003)

34. Richman J.A. and Moorman J.R., Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol.* **278**, H2039–H2049 (2000)

35. Webber, C.L. Jr. &Zbilut, J.P., Dynamical assessment of physiological systems and states using recurrence plot strategies. *J ApplPhysiol,* **76,**965–973 (1994)

36. Bracewell, R.N. The Fourier transform and its applications. Book 3rd.Ed. Boston: ***McGraw Hill*** (2000) ISBN: 0-07-303938-1

37. Vandeput, S., Verheyden, B., Aubert, A.E., Van Huffel S.M. Non-Linear Heart Rate Dynamics: Circadian Profile and Influence of Age and Gender. *Med. Eng.and Phys.* **34**, 108-117 (2012)

38. Goldberger, A.L. et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* **101**, 215-220 (2000)

39. Aliot, E., Haissaguerre,M., Jackman, W. Catheter Ablation of Atrial Fibrillation. Book ***Wiley-Blackwell Ed.*** (2008) ISBN: 978-1-4051-6349-1

40. Haissaguerre, M. and Zipes,D. Thoracic Vein Arrhythmias: Mechanisms and Treatment. Book ***Blackwell Futura Ed.*** (2004) ISBN: 978-1-4051-1888-0