UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXXIII CICLO DEL DOTTORATO DI RICERCA IN

INGEGNERIA INDUSTRIALE E DELL'INFORMAZIONE

Analysis of the circadian rhythm of cardiovascular signals and their prognostic use in decision support systems

SSD ING-INF/06 Bioingegneria Elettronica ed Informatica

DOTTORANDA
Giulia Silveri

COORDINATORE
Prof. Fulvio Babich

SUPERVISORE
Prof. Agostino Accardo

ANNO ACCADEMICO 2019/2020

Cominciate col fare il necessario,
poi ciò che è possibile e
all'improvviso vi sorprenderete a fare l'impossibile.

San Francesco di Assisi

Contents

List of Figures

List of Tables

Acronyms

ABPM Ambulatory Blood Pressure Measurement

ABPMm Mean average over 24hrs

ACC Accuracy

AG Adult Group

ANN Artificial Neural Network

ANS Autonomic Nervous System

AUC Area Under the Curve

BP Blood Pressure

CART Classification and Regression Tree

DBP Diastolic Blood Pressure

DCM Dilated Cardiomyopathy Disease

DSS Decision Support System

ECG Electrocardiogram

FD Fractal Dimension

FN False Negative

FP False Positive

H Hypertensive

HD Hyperthensive with dyslipidemia

HF High Frequency

HFn High Frequency normalized

HNS Hypertensive Non Smoker

HO Hyperthensive obese

HR Heart Rate

HRV Heart Rate Variability

HS Hypertensive Smoker

IHD Ischemic Heart Disease

KNN Kth nearest-neighbours

LDA Linear Discrimination Analysis

LF Low Frequency

LF/HF Ratio between the LF and the HF

LFn Low Frequency normalized

LVEF Left ventricular ejection fraction

MeanRR Mean RR intervals

ML Machine Learning

MLP Multilayer Perceptron

NH Normotensive

NHD Normotensive with dyslipidemia

NHn Normal/High Normal Blood Pressure

NHNS Non Hypertensive Non Smokers

NHO Normotensive obese

NHob Optimal Blood Pressure

NHS Non Hypertensive Smokers

NN50 Number of pairs of successive NNs that differ by more than 50 ms

PCA Principal Component Analysis

pNN50 Proportion of NN50 divided by total number of NNs

PRE Precision

PSD Power Spectral Density

RMSSD Root Mean Square of RR intervals Successive Differences

ROC Receiver Operator Characteristic

SBP Systolic Blood Pressure

SD1 Standard Deviation of instantaneous short-term RR interval variability

SD2 Standard Deviation of continuous long-term RR interval variability

SDNN Standard Deviation of Normal to Normal RR intervals

SEN Sensitivity

SG Senior Group

SLP Single Layer Perceptron

SPE Specificity

SVM Support Vector Machine

TN True Negative

TP True Positive

VLF Very Low Frequency

YG Young Group

Abstract

The focus of my research activity has been on the processing of cardiovascular signals in order to be able to use them as a support tool for doctors in their clinical decision making. Although the analysis of these cardiovascular signals has mainly been based on conventional techniques, involving the punctual estimation of blood pressure and heart rate parameters (such as office measurement or the single average over 24hrs), it is already well known that the outpatient information, obtained from the 24hrs ambulatory monitoring, can provide useful prognostic support. Therefore, in my research activity, I have tried to examine in detail how the values of the parameters related to blood pressure and heart rate over 24hrs change and how the relationship between them varies. Furthermore, as it is known that there are numerous cardiovascular risk factors that can alter the trend of these biological signals, I have performed a detailed analysis of the effects of each single risk factors on the circadian trend of the two signals and their relationship. Since, in recent years, new mathematical approaches have been developed for the construction of clinical decision support systems applied, in the cardiovascular field, only to the classification of single heart beats of subjects suffering from different pathologies; I have, in my research activity, developed decision support systems to identify subjects with or without cardiovascular diseases. Specifically, the pathologies examined were ischemic heart (IHD) and dilated cardiomyopathy (DCM).

The previous described problems have been addressed using linear and non-linear methods of signal processing and applying artificial intelligence algorithms. In particular, the average circadian trends of pressure and heart rate on different categories of subjects and the relationship between the two cardiovascular signals over 24hrs were obtained. In addition, the main linear and non-linear parameters were calculated from the heart rate variability signal and, finally, two machine learning techniques were developed, Artificial Neural Network (ANN) and Classification and Regression Tree (CART), applied to the previous parameters in addition to age, gender and to a specific clinical parameter obtained in a non-invasive way through echocardiography, in order to identify different pathologies. The research activity was conducted in collaboration with the Department of Geriatrics and the Department of Cardiology, Azienda Sanitaria Universitaria Giuliano Isontina (ASUGI), Trieste.

The results showed that the cardiovascular signals over 24hrs show a characteristic linear circadian rhythm divisible into four precise time intervals for the pressure signal (three intervals for the heart rate) in both normotensive and hypertensive subjects highlighting the importance of taking into account the time of day in which the signal is measured. Furthermore, the relationship between these two signals evaluated over 24hrs could be useful for understanding the control mechanism of the autonomic nervous system. The examination of the effects of risk factors on cardiovascular signals (such as smoking, obesity and dyslipidemia) has shown that each single factor modifies the physiological signals. The investigation of the influence of age and gender on cardiovascular signals also highlighted a particular circadian trend, with inversion of the trend in linear and non-linear parameters of heart rate variability in subjects over 60 years of age and a gender differentiation only during the night. Finally, the results obtained by developing decision support systems based on machine learning techniques applied to various combinations of parameters, selected through principal component analysis, stepwise regression or correlated for less than 90%, showed that the ANN technique was able to identify normal subjects and IHD patients with an accuracy of 80% and that the CART algorithm was able to classify DCM patients with an accuracy of 97%. The latter technique was also able to distinguish these two pathologies from each other and from normal subjects with an accuracy of 81%.

The results of my PhD activity highlight the importance of circadian analysis of cardiovascular signals, suggesting that particular attention should be paid to the time in which the measurements are performed providing useful information for the evaluation of the mechanisms that regulate the physiological control of the two examined signals. Furthermore, the results underline how the use of decision support systems based on machine learning techniques applied to parameters obtained in a non-invasive way from the processing of the heart rate variability is useful for diagnosing various cardiovascular diseases.

Sommario

La mia attività di ricerca si è concentrata sull'elaborazione dei segnali cardiovascolari, come strumento in grado di supportare i medici nelle decisioni cliniche. Anche se l'analisi dei segnali cardiovascolari è principalmente basata su tecniche convenzionali di stima puntuale dei parametri pressori e di frequenza cardiaca (come la misura in office o la singola media sulle 24 ore), è noto che, le informazioni ricavabili da misure ambulatoriali durante le 24 ore possono fornire un utile supporto prognostico. Pertanto, nella mia attività ho cercato di esaminare nel dettaglio, come i valori dei parametri legati al sistema pressorio e cardiaco cambiano durante le 24 ore e come si modifica la loro relazione. Inoltre essendo noto che, esistono numerosi fattori di rischio cardiovascolare che possono alterare l'andamento di questi segnali biologici, ho eseguito un'analisi dettagliata degli effetti dei singoli fattori di rischio sull'andamento circadiano dei due segnali e sulla loro relazione. Poiché negli ultimi anni sono stati sviluppati nuovi approcci matematici per la costruzione di sistemi di supporto alle decisioni cliniche applicati in campo cardiovascolare, solo alla classificazione di singoli battiti cardiaci di soggetti affetti da diverse eziologie, nella mia attività di ricerca ho sviluppato sistemi di supporto per identificare i singoli soggetti affetti o meno da patologie cardiovascolari. In particolare le patologie esaminate sono state l'ischemia cardiaca (IHD) e la cardiomiopatia dilatativa (DCM).

Le problematiche precedentemente descritte sono state affrontate utilizzando metodi lineari e non lineari di elaborazione dei segnali e applicando algoritmi di intelligenza artificiale. In particolare sono stati ricavati gli andamenti circadiani medi di pressione e di frequenza cardiaca su diverse categorie di soggetti e la relazione tra i due segnali cardiovascolari nelle 24 ore. Inoltre sono stati calcolati i principali parametri lineari e non lineari dal segnale di variabilità cardiaca ed, infine, sono state sviluppate due tecniche di machine learning, Artificial Neural Network (ANN) e Classification and Regression Tree (CART) applicate ai precedenti parametri oltre che all'età, al genere e ad uno specifico parametro clinico ricavato in modo non invasivo tramite ecocardiografia, al fine di identificare diverse eziologie. L'attività di ricerca è stata condotta in collaborazione con il Dipartimento di Geriatria e il Dipartimento di Cardiologia, Azienda Sanitaria Universitaria Giuliano Isontina (ASUGI), Trieste.

I risultati hanno evidenziato che i segnali cardiovascolari lungo le 24 ore mostrano un caratteristico ritmo circadiano lineare divisibile in quattro precisi intervalli di tempo per il segnale pressorio (tre intervalli per quello cardiaco) sia in soggetti normotesi che ipertesi evidenziando l'importanza di tenere in considerazione l'ora del giorno nella quale viene misurato il segnale. Inoltre, la relazione tra questi due segnali valutata lungo le 24 ore potrebbe essere utile per comprendere il meccanismo di controllo del sistema nervoso autonomo. L'esame degli effetti dei fattori di rischio come il fumo, l'obesità e la dislipidemia sui segnali cardiovascolari ha evidenziato che ogni singolo fattore modifica i segnali fisiologici. Anche l'indagine dell'influenza dell'età e del genere sui segnali cardiovascolari, ha evidenziato un particolare andamento circadiano con inversione del trend nei parametri lineari e non lineari della variabilità cardiaca nei soggetti di età superiore ai 60 anni e una differenza legata al genere solo durante la notte. Infine i risultati ottenuti sviluppando sistemi di supporto alle decisioni basati su tecniche di machine learning applicate a varie combinazioni di parametri, selezionati mediante analisi delle componenti principali, stepwise regression o correlati per meno del 90%, hanno evidenziato che la tecnica ANN è stata in grado di identificare soggetti normali e IHD con una precisione del 80% e che l'algoritmo CART è stato in grado di classificare i pazienti DCM con una precisione del 97%. Quest'ultima tecnica è stata inoltre in grado di distinguere queste due eziologie tra loro e da soggetti normali con una precisione dell'81%.

I risultati della mia attività di dottorato evidenziano l'importanza dell'analisi circadiana dei segnali cardiovascolari suggerendo di porre particolare attenzione all'ora nella quale eseguire le misure fornendo utili informazioni per la valutazione dei meccanismi che regolano il controllo fisiologico dei due segnali esaminati. Inoltre i risultati sottolineano come l'uso di sistemi di supporto alle decisioni basati su tecniche di machine learning applicate a parametri ricavati in modo non invasivo dall'elaborazione del segnale di variabilità cardiaca è utile per diagnosticare diverse malattie cardiovascolari.

Introduction

Our bodies are constantly communicating information about our health. This information can be captured through instruments that measure physiological parameters. In particular, in the cardiovascular system the Heart Rate (HR) and Blood Pressure (BP) have pivotal roles in both health and disease. Biomedical signal processing involves the analysis of these measurements to provide useful information upon which clinicians can make informed decisions.

Most of these biological processes occur in an appropriate temporal sequence following natural rhythms, called 'circadian rhythms', ensuring that human physiology is organized around the daily cycle of activity and sleep. The 24hrs rhythm not only dictates the endogenous sleep/wake cycle, but also influences behavior and nearly every physiological function.

Until now, from a clinical prospective, understanding the interconnections of the cardiac autonomic nerves and central commands on the cardiovascular system has been widely debated. In order to evaluate the mechanisms that regulate the circadian behavior and to accurately describe the relation between circadian rhythm of BP and HR, the quantitative description of this relationship is reported as one of the main topics of this thesis.

Moreover, to explore how the time of day influences possible changes in the relationship between BP and HR, in this thesis the hour-to-hour relationship is examined comparing the results with those obtained considering only a single measurement or the 24hrs averaged values.

Finally, from the literature, it is known that some risk factors could modify the mean values of cardiovascular signals during the day time, the night time or during 24hrs, but the effects along the 24hrs are not still accurately described. In particular, the study of the rhythm, hour-to-hour, could help to understand if the two system controlling cardiac and pressure variability could work independently and what is the influence on such rhythm of several cardiovascular risks. With this in mind, this thesis will examine how the rhythms change during 24hrs due to the presence of specific risk factors such as smoking, obesity and dyslipidemia . In addition to this, how age and gender alter the cardiovascular signals has also been described too.

Beside the analysis of cardiovascular signals, different mathematical approaches for Decision Support System (DSS) have been proposed as a useful research

diagnostic tool for physicians with the aim of early disease identification that reduces the effects of pathology and mortality rate. Different clinical DSS based on new ways to process signals, such as machine learning techniques, have been recently introduced. However in the cardiovascular field, the DSS has been developed only for classifying single heart beats or RR segments, taken from public databases, and not for identifying subjects affected or not by cardiovascular disease. In this thesis, I developed some decision support systems able to predict different cardiovascular pathologies, analyzing features extracted from a large dataset of subjects. In particular, two cardiovascular pathologies were considered: Ischemic Heart Disease (IHD) and Dilated Cardiomyopathy Disease (DCM). The DSS were applied on non-invasive features such as several parameters extracted from the time variability of the interval between heartbeats (Heart Rate Variability (HRV)) and only one clinical parameter (Left ventricular ejection fraction (LVEF)), taken from echocardiography.

In the first chapter, the state of art about BP and HR cardiovascular signals using conventional measurements such as office and the mean over 24hrs has been reported, highlighting that the knowledge about their relationship over 24hrs has never been completely studied. In addition, the literature concerning the influence of time of the day on this relationship has been reported. Since several risk factors affect the cardiovascular rhythms, the literature about the changes due to the presence of smoking, dyslipidaemia, age and gender on each signal and on their relationship has been reported. Finally, the literature about the identification of some cardiovascular pathologies, in particular IHD and DCM, based on particular machine learning techniques, has been presented. In the second chapter, methods concerning BP and HR signal processing and artificial intelligence algorithms used for pathologies identification have been described. In the third chapter, the results of my research activity, performed during the three years of my PhD program, were reported. The studies were carried out in collaboration with the Geriatric and Cardiovascular Departments, Azienda Sanitaria Universitaria Giuliano Isontina (ASUGI).

Chapter 1

State of art

In the first chapter, the cardinal vital signals, BP and HR, and their variations over a 24hrs period have been introduced. Successively, the literature concerning the relationship between these two cardiovascular circadian rhythms in a specific pathological condition, such as hypertension, has been reported. To better analyze and quantify this relationship, the effect of some cardiovascular risk factors such as age, gender, smoking, obesity and dyslipidemia on each signal behavior and on their relationships have been studied. Finally, the literature concerning the artificial intelligence techniques used to identify cardiovascular disease have been mentioned, with particular attention on decision support system for classifying Ischemic Heart and Dilated Cardiomyopathy diseases.

1.1 Cardiovascular signals

Variability signals related to the cardiovascular system contain relevant information about the behavior of the autonomic nervous system that acts as a controller of many physiological parameters such as heart rate and blood pressure [1, 2, 3, 4]. In particular, HR measurements provide significant prognostic information about cardiovascular risks [5] and BP measurements represent a powerful prognostic marker of target organ damage [6].

1.1.1 Blood Pressure

BP measurement is used to monitor cardiovascular health. It estimates the force applied to the blood vessels during blood circulating, it decreases as it moves away from the heart through arteries and capillaries and toward the heart through veins. The highest pressure occurs when blood is traveling through the arterial circulation by the contraction of the heart which is known as the Systolic Blood Pressure (SBP), while Diastolic Blood Pressure (DBP) measurement is taken when the heart relaxes between beats and the pressure in the arterial circulation falls to its lowest level.

Blood Pressure monitoring

There are several methods to measure BP, such as "office", "out of office", "home" and "ambulatory monitoring" as reported in [7]. The office measurements have been considered as the standard measurement techniques for diagnosis. The office blood pressure is most commonly measured via a sphygmomanometer which consists of a combination of cuff, inflating bulb (with a release valve) and a manometer. Instead, the Ambulatory Blood Pressure Measurement (ABPM) is a non-invasive, fully automated technique in which BP is recorded over an extended period of time, typically 24hrs, to have an accurate estimation. The device is typically programmed to record BP at 15 - 30 min intervals and average BP values are usually provided for day time, night time and 24hrs. The instrument used is a portable blood pressure machine worn as a belt, with the cuff being attached around the upper arm. This measurement is able to reduce the white coat hypertension effect in which a patient's blood pressure is elevated during the examination process due to nervousness and anxiety caused by being in a clinical setting. ABPM can also detect the reverse condition, masked hypertension, where the patient has normal blood pressure during the examination but uncontrolled BP at home. The other two techniques, out of office and home, were not considered because they do not provide useful information during the night and because they provide approximate prognostic evidence.

1.1.2 Heart Rate and Hear Rate Variability

Heart rate is the measure of the number of times per minute that the heart contracts. The speed of the heartbeat varies as a result of physical activity, threats to safety, and emotional responses. The measure of the variation in time between heart beats is called Heart Rate Variability (HRV) and it is used to quantitatively assess the cardiac autonomic activity as a result of the interaction between sympathetic and parasympathetic activity. Nowadays, an increased interest in understanding the HRV mechanism and its clinical utility in diseases has been developed. Changes in HRV patterns seem to be a sensitive indicator of health impairments, providing valuable insight into pathological conditions and enhancing risk stratification.

Heart Rate and Hear Rate Variability measurement

The office heart rate is taken in the physician's office and it refers to the HR when a person is relaxed. It provides the number of times the heart beats in the space of a minute by manually evaluating the pulses of the radial artery. The HRV signal is usually extracted from an Electrocardiogram (ECG) signal obtained from a 24hrs Holter monitoring device. Each QRS complex is detected from a continuous ECG recording in order to quantify the fluctuations in the intervals between heart beats known as RR intervals [8]. This ambulatory electrocardiograms is a battery-operated portable device. It is the size

of a small camera and it has wires with silver/silver chloride, dollar-sized electrodes that attach to the skin. To obtain indexes able to characterize HRV, the RR time series should undergo a proper mathematical processing in order to avoid incorrect conclusions and unfunded extrapolation.

1.2 Circadian rhythms

The circadian rhythms are physical, mental and behavioral changes that follow a 24hrs cycle. These natural processes are due primarily to light and dark. One example of the light-related circadian rhythm is sleeping at night and being awake during the day. Almost all the cardiovascular signals presented a circadian rhythm characterized by oscillation over 24hrs period, including BP and HR. Although, conventional office BP and HR measurements have long been used to evaluate physiological states [9, 10], they did not provide additional prognostic information about short- and long-term fluctuations during the day [11, 12, 13, 14] as estimated by using ABPM. During my PhD, I have analyzed this topic focusing on the relation between the circadian rhythm of these two biomedical signals in a specific pathological condition such as hypertensive and normotensive subjects. Successively, the knowledge about influence of the time on this relationship, as an important element on clinical outputs, has been presented. Moreover, the literature about the influence of several cardiovascular risk factors on each circadian signal and on their relationship has been reported. Finally, since ageing influences the BP increase, the influence of several cardiovascular risk factors on this relationship was examined.

1.2.1 Relation between Blood Pressure and Heart Rate circadian rhythms and their differences among hypertensive and normal blood pressure subjects

Although considerable knowledge has been gained about the connection between BP and HR, the relationship over 24hrs has never been completely described, although it is known that both signals have a circadian rhythm and they can be measured with non-invasive techniques [15]. Several authors highlighted that the HR during sleep presents lower values than during day time [16, 17, 18] and the BP shows a higher level during the day time than during the night [19, 20]. Moreover, the day–night difference in the HR was positively associated with the day–night differences in office SBP and DBP. Hence, the diurnal variations of HR tended to parallel the diurnal variations in BP and to decrease with reducing BP [21]. Koroboki et al. [20], using ABPM measurements, studied the circadian variation of HR and BPs in 1676 Greek subjects showing that patients with masked hypertension and white coat hypertension presented the same mean pattern as normotensive and hypertensive subjects. The circadian rhythms presented peak values between 10:00 and 14:00, an

afternoon decrease with a second peak at about 20:00 and a progressive decrease till 2:00–4:00. Moreover, analyzing only SBP in normotensive adults over 24hrs, Madden et al. [22] approximated, by means of piecewise linear splines, the steepest rise and fall occurred just after waking and immediately after falling asleep. In order to fill in the lack of research about the relationship between BP and HR over 24hrs in my research activity [1SG] [2SG], the differences in circadian changes among hypertensive subjects and normotensive (normal/high and optimal blood pressure) subjects was examined and quantified.

1.2.2   Influence of the time of day on the relationship between Heart Rate and Blood Pressure

The time taken to achieve BP and HR and to evaluate their relationship is an important determinant of clinical outcomes and it is still being debated. Most studies utilized conventional BP monitoring performed in the doctor's office that is the common way used to evaluate the subjects pressure health state [23, 24]. However, from a pathophysiological point of view, the correlation between the value of the BP and the target organ damage or cardiovascular risk has greater prognostic significance when the BP values are obtained by ABPM than when they are taken in office [25]. Thus, Goldenberg and al. [26] proposed to use ABPM measures evaluating the relationship between the average of both BP and HR over 24hrs in normal, isolated systolic hypertension and essential hypertension. They showed that, in isolated systolic hypertension, there was a negative relationship between SBP and HR while in the other two groups no correlation was found. On the other hand, they found a positive relationship between DBP and HR in all the groups. Furthermore, it is well known that the circadian rhythms of HR and BP signals change over 24hrs [27]. These variations could affect HR and BP signals differently and, therefore, their relationship during each hour of the day. On the other hand, the hour-to-hour study would allow a more accurate analysis of this relationship compared to a punctual one, like that used untill now; either in the office or as a simple average over 24hrs. This analysis would also allow the examination of how the inputs from cardiac autonomic nerves in response to receptors and from central autonomic commands could alter cardiovascular control system over 24hrs. In order to explore how the time of day influences possible changes in the relationship between BP and HR, in my research activity this hour-to-hour relationship was examined and compared with the results from those obtained considering office or 24hrs averaged BP values [3SG].

1.2.3 Influence of smoking on Heart Rate and Blood Pressure circadian rhythm in hypertensive and non hypertensive subjects

Hypertension and smoking are considered as independent risk factors for cardiovascular and neuroendocrine function [28, 29]. In normal subjects, Cryer et al. [30] observed an acute rise in BP and HR in smokers immediately after the beginning of smoking, persisting over 15 min after smoking one cigarette. The rise of these clinical parameters is caused by the reduction of the cardiac baroreflex sensitivity due to nicotine [31]. Al-Safi, using office measurements, investigated the correlation of smoking with BP and HR, highlighting that smokers had significantly higher BP than non-smokers while HR values were significantly higher only in male smokers [32]. Bolinder et al. [33], studying 135 normal people, found that office HR and SBP values in smokers were significantly lower than the mean values over 24hrs obtained using ABPM. In addition, they found that the mean value of HR during day time was significantly greater in smokers than in non-smokers. On the other hand, in hypertensive patients, Toumilehto et al. [34] showed that cardiovascular mortality was three to four times higher in smokers than in non-smokers even if smokers displayed similar BP and HR office values than non-smokers. Moreover, Bang et al. [35] showed that hypertensive smokers had significantly higher BP and HR than non-smokers only during day time and that hypertensive smokers presented slightly (not significant) higher office BP values than non-smokers. In addition, Verdecchia et al. [36] found that the office BP and HR in hypertensive smokers and non-smokers was nearly identical; on the contrary, the mean on 24hrs BP was significantly higher in hypertensive smokers. Furthermore, Soresen et al. [37] evaluated the impact of smoking on office and ambulatory BP and HR values on treated and non-treated hypertensive subjects, underlining in both groups significantly greater day time ABPM and HR values in smokers than in non-smokers. The HR values were significantly higher in smokers also during night time and in office conditions. However, the use of a single value (office as well as mean on day time, on night time or on 24hrs) only approximatively describes the changes due to the circadian rhythm of BP and HR occurring during 24hrs. A better temporal definition of the changes occurring in BP and HR during 24hrs can also be useful to identify, with greater accuracy, not only possible pressure peaks but also how quickly the changes occur in BP an HR . Therefore, in my research activity during the PhD program, it was evaluated whether and how the circadian HR rhythm is modified in hypertensive and non hypertensive smoker compared to non smokers [4SG]. Successively, how the BP changes over 24hrs in normotensive and hypertensive smokers compared to non-smokers was observed [5SG]. Moreover, to reduce the effects of other cardiovascular risk factors such as obesity, dyslipidemia and diabetic mellitus, which generally are positively associated with BP increasing its value, the hypertensive and normotensive

subjects, which did not present these risk factors, were considered.

1.2.4   Effects of other risk factors (obesity and dyslipidemia) on the cardiovascular circadian rhythm

There are many risk factors associated with alterations of cardiovascular signals. Some risk factors, such as age and gender, cannot be modified, while other risk factors, like obesity, dyslipidemia and smoking can be modified with treatment. However, to do that, it is important to understand the period of time over 24hrs in which the signal's alteration is manifested.

The presence of cardiovascular risk factors such as smoking, obesity and dyslipidemia affects the sympathetic activity of both normotensive and hypertensive subjects, modifying the HR circadian rhythm [32, 33, 34, 35, 36, 37, 38, 39]. In particular, obesity is considered a potential predictor for cardiovascular disease due to the impairment in the autonomic nervous system controlling HR. In normotensive subjects, Lee et al. [38] found higher values of HR in obese than in non-obese subjects in a cohort of 33 subjects highlighting that office HR was 14 beats greater in non-obese group (54 beats/min vs 40 beats/min). This punctual result was confirmed by Rossi et al. [40] that extended the finding also to the mean HR over 24h in a population of 92 individuals, underling higher mean HR values in obese subjects. This could be due to the fact that, in obese subjects, the elevated sympathetic activity affected the peripheral vessel and this alteration may be related to autonomic abnormalities [38, 40]. Grassi et al. [41] found that in 10 young obese subjects and 8 age matched eutrophic subjects there were no differences in HR between the two groups. Additionally, Junior et al. [42] in 180 obese children between 7-16 years has shown that being overweight and obesity are accompanied by higher HR values bringing to an alteration in autonomic mechanism. In addition, in a survey studying [39] of 3464 adults with hypertension, the authors underlined that obesity was associated with increased HR (75 ± 11 beats/min vs 73 ± 10 beats/min), which may at least in part reflect increased arterial stiffness and increased sympathetic tone. Considering both day and night time, Kotsi et al. [43] found a significant correlation between obesity and mean HR values in 3216 hypertensive subjects, with higher values during day time. In regards to dyslipidemia, Sun et al. [44], in a study of 9415 normotensive subjects aged >40 years old, underlined that the cardiovascular risk due to dyslipidemia is related to higher values of office HR due to an activation of the sympathetic discharge. In addition, Lee et al. [45], in a normotensive population of American Indians, underlined that high values of dyslipidemia presented higher values of office HR. In hypertensive subjects, Perlini et al. [39] found an increased office HR associated with dyslipidemia. Moreover, Palatini et al. [46], in elderly hypertensive subjects, found that HR mean values, evaluated over 24hrs, showed significant correlations with dyslipidemia. In order to separately evaluate the influence of specific risk factors (such as: obesity, dyslipidemia and smoking)

on the HR circadian rhythm, in my research activity, how the rhythm during 24hrs changes in normotensive and hypertensive subjects presenting only one risk factor at a time [6SG] was carefully examined.

On the other hand, ageing is considered as one of the natural causes affecting the increase in BP. An epidemiological study on normal subjects estimated that SBP increases with an average of 140 mmHg by the seventh decade, and that the DBP tends to increase with age but the average value tends to remain flat or to decline after the fifth decade [47]. Franklin et al. [48], in The Framingham Heart Study, characterized the age related changes in BP in a sample of 20136 normotensive and untreated hypertensive subjects (50 to 79 years old) highlighting that DBP falls after 60 years old while SBP continually rises. Moreover, the differences between office and ambulatory BP increase progressively both with age and office BP values; office measurements were also higher than ABPM values both in normotensive and hypertensive subjects [48, 49, 50]. On the contrary, in young normal subjects, aged 4-18 years old, ABPM values were most often higher than office BP values and this difference was reduced with ageing [51]. In a meta-analysis on 13 normal population base cohorts, Ishikawa et al. [52] found that BP office increased with age more steeply than ABPM only after 50 years old for SBP and after 45 years old for DBP, but office BP was lower than ABPM in the youngest. Furthermore, Conen et al. [53] compared individual differences between ABPM and office BP according to 10-year age categories in subjects not taking antihypertensive treatment, finding that office SBP increase from 117 to 149 mmHg and DBP from 64 to 82 mmHg from the youngest to oldest age. Additionally, they highlighted that the relationship between age categories and diastolic ABPM increased until 50 years old and then it decreased. They also observed that among subjects younger than 50 years old, ABPM were higher than office; for ages between 50 to 60 years the ABPM and office were similar, and, when older than 60 years, this relationship was inversed. Moreover, it is known that BP is correlated with some risk factors of cardiovascular disease such as smoking, obesity, dyslipidemia and diabetic mellitus that generally increase its value [47, 54, 55, 56, 57]. Although BP values increase due to both ageing and some risk factors, until now the studies have examined the relationship between BP and age without quantifying the influence of risk factors on this relation [47, 48, 49, 51, 52, 53]. Since, from a clinical point of view, it can be useful to understand how the relationship between BP and age is affected by risk factors in people with and without hypertension, during my PhD program it was accurately examined how the relation between age and BP changes in subjects presenting at least one risk factor. The considered factors were smoking, obesity, dyslipidemia and diabetic mellitus, in subjects whose BP were evaluated in both office and ABPM ways [7SG].

1.2.5   Influence of age and gender on cardiovascular behaviors over 24hrs in normotensive subjects

It is known that the world population is progressively ageing and to increase the health span alongside the lifespan [58], it is necessary to focus attention on specific cardiovascular predictors in the elderly. Although several reports described the analysis of HRV as a promising clinical tool in different conditions [59], to date HRV measurements are still far away from mainstream clinical practice. On the other hand, HRV shows a regular circadian rhythm controlled by the vagal system during day and night [60] but its rhythm is altered in some pathologies like acute myocardial infarction and idiopathic dilated cardiomyopathy [61, 62]. Moreover, in normal people, ageing modifies the amplitude of this rhythm with a reduction in the elderly [63]. Since a focused and comprehensive HRV analysis in a wide population of elderly subjects over 75 years old and a direct comparison with younger counterparts was still missing, in this thesis the influence of age on HRV was evaluated over 24hrs in normal subjects, from 15 to 90 years old, using both linear and non-linear parameters to identify a possible model of this relation [8SG].

As mentioned, the increases and reductions along a 24hrs period of cardiovascular signals are the result of stimulation and deactivation due to the sympathetic/parasympathetic systems. Furthermore, it is known that a gender-related difference in the baroreceptor reflex control of both BP and HR existed and that the females have a significantly smaller baroreflex sensitivity than males [64]. The lower baroreflex responsiveness makes females less able to compensate for a cardiovascular event and put them at increased risk of death [65]. Some authors [23, 66, 67], measuring BP and HR variables in office condition, highlighted that females had significantly higher values of HR and lower values of BP than males. In particular, Morcet et al. [23] in a large population showed that the punctual measurement of BP is higher in females than in males. Moreover, Zhang and Kesteloot [24], in a study of 5027 males and 4150 females, showed a significant positive association between HR and BP, with SBP more strongly correlated with HR than DBP, in both genders. On the other hand, using ABPM measure, Khoury et al. [67] pointed out that both SBP and DBP were higher in males than in females, in a cohort of 69 males and 62 females. Also Thayer et al. [68], in a sample of 33 young subjects (19 males and 14 females), confirmed higher values of mean BP in males than in females. On the contrary, Hermida et al. [69], showed higher values in females than males, underlining statistically significant differences between males and females mostly in SBP and in HR (not in DBP). Additionally, Jaquet et al. [70], found that SBP and DBP had a significant gender difference only in young groups and that elderly females had higher BPs than males; both elderly and young groups displayed gender differences for HR values higher in females than in males. Since the relationship between BP and HR, depending on gender, could be related to heart disease [6, 9] in this thesis,

how gender affects the 24hrs BP/HR relation using ambulatory measurements in normotensive subjects was examined and quantified [9SG].

1.3 Artificial intelligence for cardiovascular diseases

Impaired regulation of the cardiac autonomic nervous system is one of the most important pathophysiologic changes in patients with heart failure and HRV can provide important prognostic information [71]. On the other hand, the diagnostic role of HRV analysis in patients with either ischemic or non-ischemic heart disease is still debated [71, 72, 73] and the use of this tool in the early differentiation between ischemic heart (IHD) and dilated cardiomyopathy (DCM) diseases is largely unexplored.

Ischemic heart disease is a pathological condition characterized by an imbalance between myocardial oxygen supply and demand, mainly due to coronary artery atherosclerosis [74]. Although the narrowing can be caused by a blood clot or by constriction of the blood vessel, most often it is caused by buildup of plaque. Typically, this disease is silent, showing no symptoms, but can become unstable due to plaque rupture or erosion causing angina, myocardial infarction or sudden cardiac death [75]. Several clinical diagnostic techniques to confirm the disease, such as those based on single photon emission computed tomography, radionuclide myocardial perfusion imaging, positron emission tomography, cardiac magnetic resonance imaging and multi-slice computed tomography are suggested by current clinical guidelines to identify this pathology. All these techniques present invasiveness and at least one limitation, such as low sensitivity, specificity, high cost and/or long examination time [76]. On the other hand, the basic testing, in patients with suspected IHD, includes laboratory biochemical test, ambulatory electrocardiogram (ECG) monitoring, stress test ECG, resting ECG and resting echocardiography. Ambulatory electrocardiogram is a tool that records the electrical activity of the heart over the time and it may detect silent myocardial ischemia by examining deviations in the ST segment or/and changes in T waves that are due to the disease. These ECG abnormalities could occur during normal daily activities as well as during the night. However, the ECG alone could not be sufficient for the diagnosis in case of a slight alteration of the repolarization, due to IHD.

On the other hand, dilated cardiomyopathy is a progressive disease of the heart muscle defined by the chamber enlargement and contractile dysfunction of the left ventricle in the absence of chronic pressure and or/volume overload; this disease is the third most common cause of heart failure [77]. If no cause is discovered the cardiomyopathy may be referred as idiopathic, characterized by an advanced stage of left ventricular dilatation and dysfunction. The DCM could be diagnosed in asymptomatic individuals during routine medical screening in

which the electrocardiogram presented abnormalities ranging from isolated T wave changes to septal Q waves. In this patient population, increased QT variability has been independently associated with occurrence of major arrhythmic events, including sudden cardiac death [78]. In literature, some studies used HRV measurement, extracted by ECG [79], as an instrument to provide additional valuable insight into physiological and pathological conditions, both to enhance risk stratification and to be a predictor of the risk of mortality [80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90]. Finally, from literature [72, 75], it has been highlighted that the left ventricular ejection fraction (LVEF) parameter, taken from the echocardiography, has been accepted as a prognostic indicator of patients with cardiovascular diseases. In particular, it was a basic non-invasive clinical tool to support the diagnosis of IHD and DCM.

In that clinical field, machine learning algorithms have been developed using different mathematical methods and clinical parameters. In this section, a literature review has been reported to understand the evaluation of these algorithms in classifying IHD and DCM patients.

1.3.1 ANN for Ischemic Heart Disease

Some authors [72, 73, 91] have underlined that the time variation of the interval between consecutive normal heartbeats (HRV) due to the interaction between sympathetic and parasympathetic activities in autonomic functioning [92], presented significant differences between IHD and normal subjects. In particular, using linear and non-linear features, they highlighted in IHD patients a decrease of the HRV parameters values in the time domain [72, 91] as well as of the normalized low frequency power, of the ratio between low and high frequency powers [91]. Moreover, Voss et al. [73] suggested that the SD1 parameter, extracted from the Poincaré plot, is a prognostic index to differentiate IHD from normal subjects. Nowadays, different mathematical approaches for DSS, based on HRV [93, 94, 95, 96] as well as on some clinical features [97, 98], have been developed to identify beats and arrhythmias and to classify cardiovascular diseases like atrial fibrillation, left bundle branch block, cardiomyopathy and ventricular fibrillation. These DSS are based on widely applied linear programming classification methods like the Kth nearest-neighbours (KNN) [93], clustering [94], Linear Discrimination Analysis (LDA) [95, 99], fuzzy analysis [100, 101, 102], Classification and Regression Tree (CART) [103, 104], Support Vector Machine (SVM) [105, 106] and Artificial Neural Network (ANN) [107, 108, 96, 97]. Some of these algorithms have also been used to classify subjects suffering from IHD. In particular, Kannathal et al. [102] used an adaptive neuro-fuzzy network applied to three non-linear parameters extracted from HRV (Largest Lyapunov Exponent, Spectral Entropy, Poincaré plot) obtaining an accuracy of 96%. However, the data were taken from half-hour segments of the 47 subjects belonging to MIT-BIH arrhythmia

database[1] of which only a few subjects presented ischemic/dilated cardiomyopathy. Acharya et al. [108], using the same database[1] and the same non-linear HRV parameters, developed a cardiac abnormalities classification system by using an ANN, obtaining an accuracy of 83.3% for identification of patients with ischemic/dilated cardiomyopathy from normal subjects. Finally, Dua et al. [96] extracted non-linear features from HRV (by using recurrence plots, Poincaré plot, detrended fluctuation analysis, Shannon's entropy, approximation entropy and sample entropy) and, after the application of main component analysis, used the first six most significant principal components and different classification techniques to classify 143 samples recorded on 10 IHD and 10 normal subjects. Using ANN with 4 hidden layers, they obtained a classification accuracy of 89.5% on this small sample of subjects. Conversely, some authors applied machine learning techniques to identify IHD patients without considering HRV parameters among their inputs. In particular, Rajeswari et al. [97], using an ANN, selected 12 clinical features able to identify 712 IHD patients. They selected age, gender, menopause, body mass index, waist circumference, systolic and diastolic blood pressure, diabetes, cholesterol, hereditary, personal habits and stress features from 17 initial parameters, obtaining an accuracy of 82.2%. In addition, Kukar et al. [98], by using some machine learning techniques such as the Bayesian classifier, artificial neural network, decision tree and k-nearest neighbor method, evaluated the performance of some diagnostic methods like clinical examinations, exercise ECG testing and myocardial scintigraphy for IHD identification in order to reduce the number of subjects to be submitted to coronary angiography. They used a dataset of 327 IHD patients, obtaining with a multilayer feedforward neural network, an accuracy of 92%, reducing the number of subjects unnecessarily submitted to much more invasive and dangerous examination by 12.2%. Although several linear and non-linear HRV parameters and non-invasive techniques individually proved to achieve a good performance in classifying IHD beats or tracts, until now the studies using ANN or other machine learning methods have been carried out only on a very limited number of IHD patients and separately using either linear or non-linear HRV parameters. On the other hand, although the correct procedure to examine HRV signal involves a pre-processing of the ECG by detecting the ectopic beats, providing a clean normal-to-normal inter-beat series containing information about the control systems that govern sinus node stimulation, the identification of ectopic beats represents a difficult task since these beats can have a waveform similar to the normal ones. Moreover, since in the IHD patients the ectopic beats occur more frequently than in the normal heart probably providing independent information on ischemia, in this thesis the heart rate total variability was evaluated, considering RR segments containing both normal and ectopic beats. However, it is worth considering that the HRV signal contains two main components of different origins, leading to opposite information about the ischemic/normal heart. The first component is

---

[1]www.physionet.org

generated by the control of the Autonomic Nervous System (ANS) on the sinus node and generally decreases in conditions of ischemic heart disease compared to normal subjects. The second component is due to ectopic beats that mostly affects the ischemic heart and increases in the case of ischemic heart disease compared to normal. Therefore, it is not easy to link an increase of one of these variability indexes to a normal or ischemic heart. However, these two components are of different types and affect HRV parameters differently. In fact, typically ectopic beats lead to a substantial increase in Root Mean Square of RR intervals Successive Differences (RMSSD), but a generic flat increase in spectral power density. Instead, a deep respiratory sinus arrhythmia (characteristic of a healthy ANS and heart, producing series of RR intervals with quite large beat-to-beat interval variations) leads to a generic increase in RMSSD but to a specific increase of the high frequency power. Therefore, it may be interested to evaluate whether an ANN using, as input, a set of parameters extracted directly from the RR series, without pre-processing for the exclusion of ectopic beats, is able to identify patients with IHD. Thus, in my PhD research the accuracy of several multi-layer feed forward neural networks based on linear or/and non-linear parameters extracted from HRV together with age and gender was assessed. In order to evaluate which combination of them produces the best performance for the identification of IHD patients, the results were validated on a large sample of subjects. In addition, the left ventricular ejection fraction, a non-invasive clinical parameter, was added and examined if it could improve the classification performance [10SG] [11SG] [12SG].

1.3.2   CART for Dilated Cardiomyopathy

Linear (time and spectral) and non-linear features of HRV have been evaluated for detection of cardiovascular disease like heart failure. In particular, depressed HRV parameters indicated an impairment of the ANS that was observed for DCM [82, 83, 84, 85, 86, 87, 88, 89, 90]. The time and frequency HRV parameters had higher values in normal than in DCM [83, 84, 85, 86], except for High Frequency normalized (HFn), that showed higher values in DCM patients [84]. However, due to the non-linear nature of heart signals, the features over time and spectral domain analysis do not characterize enough subjects suffering from cardiovascular disease. Moreover, as the noise in signal increases, the effectiveness of spectral domain analysis will decrease, so some authors introduced non-linear features extracted from HRV signal like sample entropy and Poincarè parameters that presented lower values in DCM [87]. Nowadays, to support decision-making, many machine learning-based methods are widely applied in the field of medicine to solve clinical problems as well as to reduce time of pre-screening process and some of these algorithms have been used to classify cardiomyopathy [87, 93, 94, 95, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108]. In particular, Adetiba et al. [87] developed an automated heart detection model using 400 ECG data segments recorded in 40 athletes and ANNs in

order to differentiate normal and cardiomyopathy heart conditions. Some statistical parameters (mean, median, mode, variance and standard deviation) calculated on each data segment were used as inputs of the ANNs with 10 hidden layers neurons, achieving an accuracy of 98%. Megat Ali et al. [88] presented a cardiomyopathy detection approach using a multilayer perceptron network based on the width of P-wave, QRS-complex and T-wave characteristics of 600 beats, from PTB Diagnostic ECG Database[1], achieving an accuracy of 98.9%. Ghosh et al. [89] developed support vector machine (SVM) to classify normal and cardiomyopathy segments using ECG signal, recorded for few minutes on 5 normal subjects and 5 patients taken from PTB Diagnostic ECG Database[1]. The SVM was used to classify the features extracted through a continuous wavelet transform of the ECG signals, obtaining an accuracy of 92%. Shukuri et al. [90] proposed the use of recurrent neural network for detecting cardiomyopathy in ECG beat samples, 200 waveforms from 52 normal and 200 waveforms from 18 cardiomyopathy subjects, taken from PTB Diagnostic ECG Database[1], too. They used five hidden neurons and four different learning algorithms, obtaining an accuracy of 90%. As these studies focused their attention on the characteristics of single beats that could be strongly affected by noise or artifacts and since cardiomyopathy also produces changes in the variation over time of the interval between consecutive heartbeats, other authors [109, 110] turned their attention toward parameters extracted from HRV signals. In particular, Thirugnaman et al. [109] studied 25 linear and non-linear HRV parameters calculated on ECG samples, 16 from DCM and 6 from normal subjects, taken from Physionet database[1]. They developed a stacked-ensemble classifier using SVM, KNN and decision tree to identify cardiomyopathy, achieving a sample detection accuracy of 99.9%. Moreover, Mahesh et al. [110] used a combination of tree structure and logistic regression on 15 linear and non-linear HRV parameters to identify 13 cardiomyopathy and 410 normal segments extracted from Physionet database[1]. Considering separately different groups of parameters, they achieved a maximum accuracy of 95.61%. Although these studies reached interesting results, none of them produced a system capable of classifying subjects as rather single heart beats or segments belonging to some normal subject or with DCM. Furthermore, data coming from a very limited number of DCM subjects were examined and a combination of several mathematical models were developed. In order to focus mainly on the identification of subjects with cardiac disease rather than on single beats, in my PhD research, it was verified if HRV derived parameters applied to a simple CART [111, 112] could be able to distinguish DCM patients from normal subjects in large cohort of cases. With the aim to evaluate which features, applied to CART, could discriminate DCM patients with high accuracy, some combinations of linear and non-linear parameters extracted from HRV together with age and gender were examined. To reduce the number of features preserving the variance, principal component analysis and stepwise regression were used too. Additionally, adding LVEF, which is a prognostic

indicator in DCM patients [77, 78], was examined to improve the classification performance [13SG]. Finally, in my research activity, it was verified if HRV parameters, age, gender and LVEF applied to a CART would be able to distinguish from IHD and DCM patients from normal subjects and between them, in a large cohort of cases [14SG].

Chapter 2

Material and Methods

2.1 Subjects classification

To evaluate the circadian rhythm of physiological variables the subjects were enrolled at the Department of Medical, Surgical and Health Care, CS of Geriatrics, University of Trieste and ASUGI Hospital between October 2017 and July 2018. The subjects with no clinical or laboratory evidence of secondary arterial hypertension, absence of clinical evidence of hypertension-related complications, no cardiac disease, no patients with type 1 diabetes were considered.

Based on office BP readings, subjects were classified as Hypertensive (H) (SBP $>$ =140mmHg and/or DBP $>$ =90mmHg) or Normotensive (NH) (SBP $<$ =140mmHg and DBP $<$ =90mmHg), according to current Guidelines [7]. Successively, NH subjects were further sub-classified as having Normal/High Normal Blood Pressure (NHn) (SBP 120–139mmHg and DBP 80–89mmHg) or Optimal Blood Pressure (NHob) (SBP $<$ 120mmHg and DBP $<$ 80mmHg).

On the other hand, ischemic (IHD) and non-ischemic (DCM) heart disease patients as well as Normal subjects afferent at the Cardiovascular Department, University of Trieste and ASUGI Hospital between September 2019 and December 2019 were enrolled. The patients were identified following the guidelines [74, 75, 77, 78]. In particular, to identify IHD patients, typical symptoms such as angina, laboratory test (increase of Troponin I), ECG repolarization abnormalities (T-wave and ST), wall motion abnormalities (hypokinesia or akinesia), detected with echocardiography or positive stress testing, were examined. Finally, the diagnosis was confirmed by coronary angiography result [76]. Following protocol, a stress test ECG was performed only in patients with intermediate CAD risk. The IHD patients with chronic coronary syndromes were considered excluding patients with the left ventricular ejection fraction lower than 50%, in such way the patients did not present systolic heart failure. The DCM disease is characterized by left ventricular dilatation and left ventricular systolic dysfunction with normal wall thickness in the absence of abnormal loading conditions [77]. This cardiovascular disease is defined by

echocardiography as the presence of an ejection fraction below 45% and/or a fractional shortening less than 25% and a left ventricular end-diastolic dimension greater than 112% of the predicted value corrected for age and body surface area. Finally, normal subjects did not present neither peripheral artery disease, thyroid disorders, history of myocardial revascularization, hypertensive heart disease, pulmonary hypertension or severe valvulopathy.

For each study, the description of the subject groups are reported in Table 2.1.

Section	Article	Total subjects	Groups of subjects
Relationship between BP and HR circadian rhythms and their differences among hypertensive and normal blood pressure subjects	1SG	629 subjects	423 Normotensive, 205 Hypertensive
	2SG	385 subjects, matching by age	215 Hypertensive 112 Normal/high normal blood pressure 58 Optimal Blood pressure
Influence of the time of day on the relationship between heart rate and blood pressure	3SG	388 subjects	216 Hyperthensive 172 Non Hyperthensive
Influence of smoking on HR and BP circadian rhythm in hypertensive and non-hypertensive subjects	4SG	614 subjects	58 Hypertensive Smoker 351 Hypertensive Non Smoker 39 Non Hypertensive Smokers 166 Non Hypertensive Non Smokers
	5SG	248 subjects without risk factors (obesity, dyslipidemia, diabetic mellitus)	32 Hypertensive Subjects 113 Hypertensive Non Smokers 20 Non Hypertensive Smokers 83 Non Hypertensive Non Smokers
Influence of smoking and other cardiovascular risk factors on Heart Rate circadian rhythm in normotensive and hypertensive subjects	6SG	618 subjects	83 Normotensive without risk factors 20 Normotensive Smoker 44 Normotensive with Dyslipidemia 23 Normotensive Obese 169 Hypertensive without risk factors 32 Hypertensive Smoker 99 Hyperthensive with Dyslipidemia 53 Hyperthensive Obese
Influence of some cardiovascular risk factors on the relationship between age and blood pressure	7SG	880 subjects	Considering the risk factors: 253 Hyperthensive evaluated in office 241 Hyperthensive in ambulatory 112 Normotensive evaluated in office 124 Normotensive in ABPM. Without risk factors: 54 Hyperthensive evaluated in office 60 Hyperthensive in ambulatory 105 Normotensive in office 99 Normotensive in ABPM
Influence of ageing on circadian rhythm of HRV in normal subjects	8SG	149 subjects	47 Young Group (15-39 years old) 47 Adult Group (40-64 years old) 46 Senior Group (65-90 years old)
Influence of the gender on the relationship between heart rate and blood pressure	9SG	172 subjects	50 males, 122 female
ANN for Ischemic Heart Disease	10SG	243 subjects	156 Normal 87 Ischemic Heart Disease
	11SG	965 subjects	681 Normal 284 Ischemic Heart Disease
	12SG	496 subjects, matched by gender and age	251 Normal 254 Ischemic Heart Disease
CART for Dilated Cardiomyopathy	13SG	972 subjects	773 Normal 199 Dilated Cardiomyopathy
	14SG	1133 subjects	689 Normal 263 Ischemic Heart Disease 181 Dilated Cardiomyopathy

Table 2.1: Subject groups analyzed during the PhD program.

2.2   Blood Pressure acquisition and analysis

The blood pressure was first measured in office condition, as the average of two consecutive measurements [7]. Trained personnel, using a standard mercury sphygmomanometer with a cuff of appropriate size, evaluated the BP values. In all studies the BP was measured in office condition around at 9:00, as the average of two consecutive readings and then in ambulatory way. The ABPM was carried out by using a Holter Blood Pressure Monitor (Mobil-O-Graph® NG, IEM gmbh Stolberg, Germany), based on oscillometric technique. The portable monitor was programmed to obtain ambulatory blood pressure and heart rate readings each 15-min interval throughout the day time (6:00 to 22:00) and each 30-min interval throughout the night time (22:00 to 6:00). The BP and HR values of the different subjects were aligned using common start time (10:00) since recordings could start at different times of the day (between 8:00 and 11:00). The acquisition rate was of 15 minutes throughout the day and of 30 minutes throughout the night. The measurement procedures were in accordance with the institutional guidelines and all the subjects gave their informed consent. No patient received additional medication that might affect the circadian blood pressure or heart rate rhythmicity. Data were analyzed by using a proprietary software developed in Matlab® (MathWorks, USA).

The circadian trend of the mean values of SBP and DBP among the subjects was separately examined for each subject group analyzing, in each period, the mean and standard deviation and the difference between the maximum and the minimum BP values. Since the subject groups were independent and showed a non-Gaussian distribution, the significance of the differences between groups was evaluated by the Wilcoxon rank sum test. A p-value of 0.05 was used as the level of statistical significance. Since the BP profiles during 24hrs showed to be bimodal with two minima and two maxima, the 24hrs were divided in four intervals corresponding approximately to 10:00-14:30, 14:30-19:00, 19:00-2:00 and 5:00-10:00. In each period the quite linear trend was fitted by a regression line and the linear approximation reliability was measured by using R-square statistic.

2.3   Heart Rate acquisition and analysis

ECG was acquired by using a 24hrs Holter monitor using a three channel tracking recorder (Sorin Group, Italy) that sampled the signal at 200Hz. Segments of 300s each were examined, since recordings could start at different times of the day (between 8:00 and 11:00), a common starting time, fixed 10:00, was selected for all the RR series. The segment time length has been chosen equal to 300s to have reliable values for short duration intervals and to ensure a frequency resolution of 0.01 Hz. The data were analysed by using a proprietary Matlab® (MathWorks, USA) program. Since the estimation of

the HRV indexes is affected by missing data, noise, arrhythmic events and ectopic beats, which all alter the signal and therefore compromise the reliability of the obtained indexes and degrade the clinical utility of measurements, the pre-processing of RR time series is needed in order to avoid it [3]. In particular, ectopic beats, defined as short RR followed by compensatory pause, are considered physiological artifacts and processed as a source of error in the HRV measures providing a clean normal-to-normal inert beat series. However, since in the diseases analyzed the ectopic beats are characteristic elements of the pathologies, which could have more powerful prognostic information than the RR time series without premature beats, in my research activity the RR time series including both normal and ectopic beats were considered analysing the heart rate "total" variability. However, in the thesis, the term HRV has been used also to indicate the "total" variability. Since RR series are formed by intervals that occur at non equidistant sampling times, in the pre-processing phase the time series should be resampled in order to perform analysis that require a time series having a constant sampling time. Thus, RR intervals were interpolated by cubic spline, starting from those classified by SyneScope® software as normal or ectopic, substituting beats classified as artifacts. In particular, methods based on spline interpolation are widely used because they are simple to implement and because they present a reduced-low pass filtering effect, which is nearly eliminated by increasing the order of the spline [113]. Even if complicated resampling schemes have been proposed [114], the error this editing introduces in the evaluation of the Power Spectral Density (PSD) has not been adequately documented [115, 116]. In my research, the interpolated signal was resampled at a frequency of 2Hz [8]. Successively, segments of 300s along 24hrs were individually examined and then mediated considering at least seven hours during day time and seven hours during night time to take into account the influence of day-night alteration. To minimize the effects of artefacts substitution, the RR segments were discarded if the total duration of artefacts in each segment was more than 5% of the segment duration, or the longest artefact tract was longer than 10s [117].

To characterize the circadian HR rhythm, since no large differences were found in the trend between 10:00-14:30 and 14:30-19:00, the 24hrs were divided in only three periods of time, from late morning to evening (10:00-20:00), from evening to night (20:00-04:00) and from early morning to late morning (05:00-10:00). The circadian trends of the mean value of HR among the subjects were separately examined for each subject group analyzing, in each period, the mean and standard deviation. Successively, assumed that the subject groups were independent and showed a non-Gaussian distribution, the significance of the differences between groups was evaluated by the Wilcoxon rank sum test. A p-value of 0.05 was used as the level of statistical significance. Successively, in each period the quite linear trend was fitted by a regression line and the linear approximation reliability was measured by using R-square statistic. To quantitatively describe HRV patterns, a large panel of HRV analysis methods

have been proposed, derived from both classical linear theory and non-linear dynamics [8].

Linear analysis

Linear methods can be divided into two main categories: time and frequency domain methods.

Time domain analysis

The time domain analysis is the simplest method to perform because either the heart rate at any point in time or the intervals between successive betas are determined. In a continuous ECG record, each QRS complex is detected and all intervals between adjacent QRS complexes, resulting from sinus node depolarization, are determined. In addition, more complex statistical time-domain measurements can be calculated derived both from direct measurements of the RR intervals or instantaneous heart rate, and from the differences between RR intervals. The most common parameters were:

- Mean RR intervals (MeanRR);

- Standard Deviation of Normal to Normal RR intervals (SDNN);

- Root mean square of RR intervals successive differences (RMSSD);

- Number of pairs of successive NNs that differ by more than 50 ms (NN50);

- Proportion of NN50 divided by total number of NNs (pNN50).

Frequency domain analysis

For the analysis of tachograms, various spectral methods have been applied but only the PSD analysis provides the basic information of how power distributes as a function of frequency. In my research activity, the PSD of a given time sequence was estimated by using the Welch's method [118]. Given $x(1), x(2)...x(N)$ the original data sequence that should be divided into d = 1; 2;...;L intervals, each one originates a M-points sub-sequence $x_d(1), x_d(2)... x_d(M)$. According to Welch, the PSD of each $x_d(n)$ sub-sequence is given by:

$$P_d(f) = \frac{1}{MU} | \sum_{n=0}^{M-1} (x_d(n) w(n) exp^{(-jfn2\pi)}|^2 \tag{2.1}$$

where U is the normalization factor for the power in the window function, selected as:

$$U = \frac{1}{M} \sum_{n=0}^{M-1} |w(n)|^2 \tag{2.2}$$

The Welch's PSD of the original sequence $x_d(n)$ is the average over these modified periodograms that is:

$$P_{Welch}(f) = \frac{1}{L} \sum_{d=0}^{L-1} (P_d(f)) \qquad (2.3)$$

The variance of the estimated PSD is determined through to a single periodogram estimation of the entire data record was decreased by the averaging of windowed periodograms. Using the Welch method, the segments were not overlapping and the variance of the averaged modified periodogram was inversely proportional to the number of segments used (L). The Hamming window was applied because it reduced the importance (weight) given to the outer samples of the segments (the samples that overlap) and it decreased the redundant information introduced by overlap between segments [119]. This spectral analysis decomposes the HRV signal into the subsequent characteristic components:

- High Frequency (HF) $[ms^2]$, ranging from 0.15Hz to 0.4Hz, it is considered an indicator of the activity of the vagus nerve on the heart;

- High Frequency normalized (HFn), it is used to minimize the effects of changes due to the acquisition setup and the normalization is obtained dividing the HF power by the total power spectrum;

- Low Frequency (LF) $[ms^2]$, ranging from 0.04Hz to 0.15Hz, which is due to the joint action of the vagal and sympathetic components on the heart, with a predominance of the sympathetic ones;

- Low Frequency normalized (LFn), the normalization is obtained dividing the LF power by the total power spectrum;

- Very Low Frequency (VLF) $[ms^2]$, ranging from 0.0033Hz to 0.04Hz, its physiological interpretation is not clear;

- Ratio between the LF and the HF (LF/HF), which reflected the absolute and relative changes between the sympathetic and parasympathetic components of the autonomous nervous system, by characterizing the sympathetic-vagal balance [120].

Non-linear analysis

Non-linear phenomena are also involved in the genesis of HRV and are determined by complex interactions of hemodynamic, electrophysiological and humoral variables, as well as by the autonomic and central nervous regulations. These non-linear mechanisms estimated the quality, scaling, and correlation properties of the signals. In other words, they are related to the complexity of the signal. The parameters used to measure non-linear properties of HRV

included Poincarè's plot, Beta exponent and Fractal Dimension (FD). The Poincarè plot was based on the theory of dynamical systems. The Beta exponent and FD quantify the self affinity characteristics of the signal analyzed in the frequency and in the time domain, respectively.

Poincarè plot

The Poincarè plot is a geometrical technique that describes the dynamics of heart rate variability representing of the values of each RR intervals into a simplified phase space that describes the system's evolution [121]. The Poincare plot is obtained by plotting each point of the time series, $X(n)$, against the next one, $X(n+1)$. The plot provides summary information as well as detailed beat-to-beat information on the behavior of the heart. It typically appears as an elongated cloud of points oriented along the line of identity: the dispersion of points perpendicular to the line of identity reflects the level of short-term variability, while the dispersion of points along the line of identity is thought to indicate the level of long-term variability [122]. To characterize the shape of the plot, a mathematical method is to fit an ellipse to the shape of the Poincarè plot [123, 124]. A set of axis oriented with the line of identity is then defined [125]. The axis of the Poincarè plot are related to the new set of axis by a rotation of $\theta = \pi/(4\text{rad})$ that is:

$$\begin{bmatrix} x1 \\ x2 \end{bmatrix} = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} \begin{bmatrix} X(n) \\ X(n+1) \end{bmatrix} \tag{2.4}$$

In the reference system of the new axes, the dispersion of the points around the $x1$ axis is measured by the Standard Deviation of instantaneous short-term RR interval variability (SD1), whereas the standard deviation around the $x2$ axis is measured by Standard Deviation of continuous long-term RR interval variability (SD2) (Figure 2.1). Finally, SD1/SD2 represents the relationship between these components, which is the ratio of short interval variation to long interval variation [79].
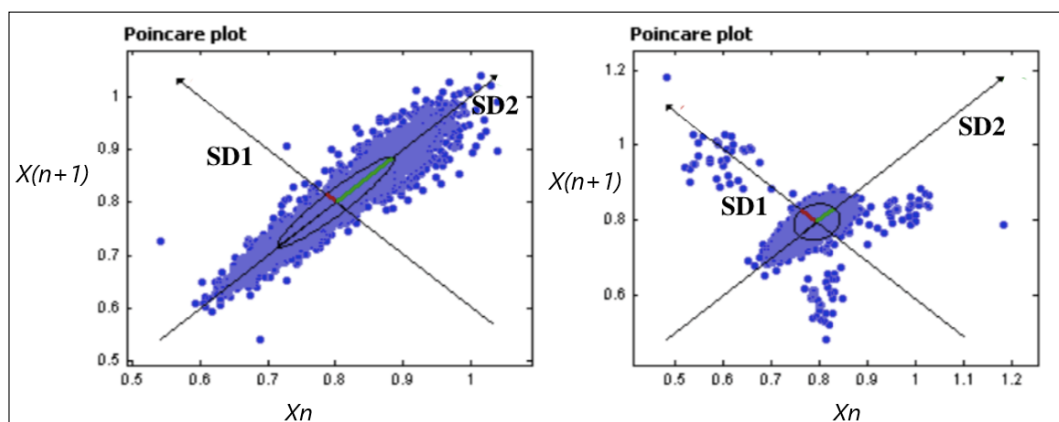

Figure 2.1: Poincarè plot for normal subject (left panel),Poincarè plot for IHD subject (right panel).

Fractal geometry of HRV

The central concept of fractal geometry is of self-similarity. This characteristic is also called scaling or scale-invariance and it describes objects that have details at a certain scale that are similar. To better explain this concept, Mandelbrot [126] introduced the term fractal. A fractal is a set of points that, when looked at smaller scales, resembles the whole set and the self-similarity is a characteristic of the fractal. This means that its details at a certain scale are similar, but not necessarily identical, to those of the structure seen at larger or smaller scales. Moreover, the notion of FD refers to a non integer dimension that originates from fractal geometry, so that the FD provides a measure of how much space an object occupies between Euclidean dimensions. From a practical point of view, the higher the FD, the more irregular is the signal. Although FD is generally meant in space, self-similar behavior can be observed also in time. The propriety of fractal time series that requires a scaling in the x and y axis, in order to appreciate their self similarity, is known as self-affinity. In a random process, the correlation between successive points is defined as the "color" of the noise, which is related to the slope of the power spectral density. The most famous is white noise, characterized by flat power spectral density. This means that, this noise is an uncorrelated process produced by a random number generator. On the other hand, the brown noise or Brownian motion, is characterized by a power density proportional to $1/f^2$. In the literature, the notion $1/f$-noise is often used to refer to any process characterized by a power spectral density proportional to $1/f^{\beta}$ (also called $1/f$-like noise).

Fractal dimension

To estimate the FD directly in the time- domain, many methods have been recommended and some comparison studies [127, 128, 129, 130] have demonstrated that Higuchi's algorithm [131] is the most accurate in estimating the fractal dimension of waveforms. It is based on the measurement of the mean length of the curve $L(k)$ by using a segment of k samples as a unit of measure. In this thesis, this method has been applied and the algorithm is reported below. Let $x(1), x(2), x(3) \ldots x(N)$ the original time series, the Higuchi's algorithm constructs k new time series:

$$x_k^m = \{x(m), x(m+k), x(m+2k), ..., x(m + \lfloor(\frac{N-m}{k})\rfloor k)\} \qquad (2.5)$$

where m is the initial time and k, ranging from 1 to k_{max}, is the discrete time interval between points. The symbol $\lfloor a \rfloor$ denotes the integer part of a. Then, for each sequence x_k^m, the length $L_m(k)$ is calculated as:

$$L_m(k) = \left[\left(\sum_{i=1}^{\lfloor \frac{N-m}{k} \rfloor} |x(m+ik) - x(m+(i-1)k|\right) \frac{N-1}{k \lfloor \frac{N-m}{k} \rfloor} \right] \frac{1}{k} \qquad (2.6)$$

Where N is the total number of samples and $(N-1)/(N-m)/k$ is a normalization factor. The length of the curve $L(k)$ for the time interval k is computed as the average of the k lengths $L_m(k)$ for $m = 1,2,...,k$. The procedure is repeated for each k ranging from 1 to L_{max}, in my research activity, it was considered k_{max}=6 and N>125 (in this PhD thesis N=600) in order to gain the reliable estimate of FD [127]. If $L(k)$ is proportional to k^{-FD}, the time series x_k^m is fractal with dimension FD. Thus, if $L(k)$ is plotted against $1/k$ on a double logarithmic scale, the data should fall on a straight line with a slope equal to –FD. The FD value is estimated by means of a least squares linear fitting procedure applied to the series of pairs $(L(k), k)$.

Power-law beta exponent

The major component of HRV occurs at frequencies below the low frequency range ($<$0.04Hz). It is known that in the range 10^{-4}-10^{-2} Hz, the power spectrum of HRV does not exhibit periodic components but a so called power law behavior. The power exponent β is relating to the PSD and to the frequency of HRV [132] as:

$$PSD \propto \frac{1}{frequency^\beta} \tag{2.7}$$

The β can be graphically visualized in a double logarithmic plot of $PSD(f)$ versus f, in which the spectrum follows a line with a slope corresponding to β. As just mentioned, the exponent β is related to the 'color' of the time series, i.e. to the degree that the series is auto correlated. If β=0, there is no autocorrelation and the PSD is flat (white noise). If β=2, then the series is highly auto correlated (brown noise). If β=1, the series is moderately auto correlated. Since, the power-law beta exponent of the $1/f$ power spectrum and the fractal dimension measure the same feature, a certain relationship is expected to exist between them, so it may also be possible to estimate one index from the other [133]. To correctly characterize HRV by either FD or β indices it is necessary to have a model that accurately describes the considered process as [131]:

$$FD = \frac{5 - \beta}{2} \tag{2.8}$$

where β is included in the 1-3 interval and FD spans the range 1-2.

2.4 Machine learning algorithms

Machine Learning (ML) techniques are a branch of Artificial Intelligence able to understand the structure of data and fit them into models. In traditional computing, algorithms are sets of explicitly programmed instructions used by computers to calculate or problem solve. Instead, machine learning algorithms develop their own models to analyze data inputs and use statistical analysis in order to output values that fall within a specific range. Because of this, machine learning facilitates computers in building models from sample data

in order to automate decision-making processes based on data inputs. In ML, tasks are generally classified into categories that are based on how learning is received or how feedback on the learning is given to the system developed. Two of the most widely adopted machine learning methods are supervised learning, which trains algorithms based on example input and output data that is labeled by humans, and unsupervised learning which provides the algorithm with no labeled data in order to allow it to find structure within its input data. In supervised learning algorithms, the individual data points in the dataset have a class or label assigned to them. This means that the machine learning model can learn to distinguish which features are correlated with a given class. Unsupervised learning involves creating a model that is able to extract patterns from unlabeled data. In other words, the computer analyzes the input features and determines what the most important features and patterns are. In my research activity, the supervised learning techniques were developed writing proprietary Matlab® (MathWorks, USA) code.

2.4.1   Artificial Neural Network

ANN is a machine learning technique designed to simulate the way the human brain analyzes and processes information [134]. The human brain consists of a large number of neural cells that work like a simple processor. A single neuron consists of three main components: dendrites that channel input signals, cell body which accumulates the weighted input signals and processes them, and axon which transmits the output signal to other neurons that are connected to it (Figure 2.2).
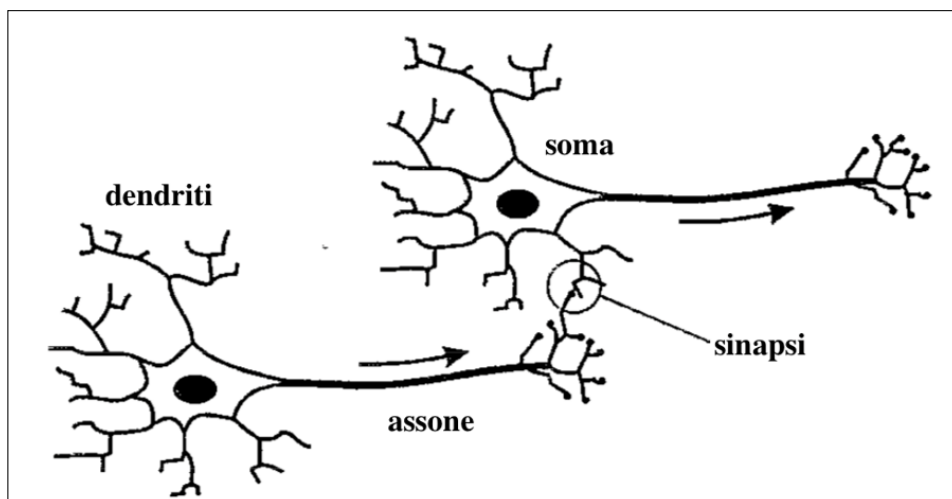
Figure 2.2: Neurons connected by synapses.

Like a biological neuron, the ANN consists of three main components: a set of input connections brings in activation from other neurons, a processing unit

sums the inputs and a non-linear activation function is applied, finally an output line transmits the results to other neurons.
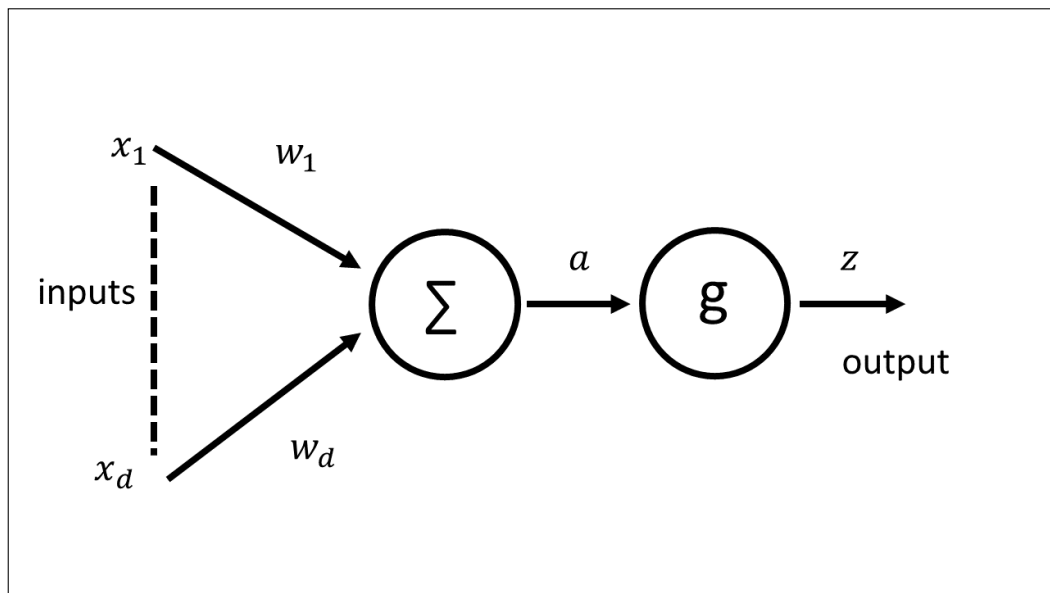
Figure 2.3: Neuron model.

The neuron is the fundamental processing element of a neural network and its simple mathematical model was introduced by McCulloch and Pittis [135]. This model can be regarded as a non-linear function which transform a set of input variables x_i ($i=1\ldots d$) into the output. The signal x_i at input i is first multiplied by a parameter w_i known as weight. The weight is analogous to the synaptic strength in a biological network and successively is added to all the other weighted input signals to give a total input to the unit. If the weights are positive indicate reinforcement otherwise represent inhibition.

$$a = \sum_{i=1}^{d} w_i x_i + w_0 \tag{2.9}$$

where the offset parameter w_0 is called bias and correspond to the firing threshold in a biological neuron. The bias can be considered as a special case of a weight from an extra input whose $x_0 = 0$

$$a = \sum_{i=1}^{d} w_i x_i \tag{2.10}$$

Finally, the output z of the unit, which may loosely be regarded as analogous to the average firing rate of a neuron, is then given by operating with a non-linear activation function $g()$ so that:

$$z = g(a) \tag{2.11}$$

The non-linearity means that the output of the function varies non-linearly with the input. The activation function does the final mapping of the activation of the output neurons into the network outputs. There are three different activation functions: threshold function, piecewise linear function and sigmoidal function. In this thesis, the non-linear curved S–shape function known as sigmoid function was studied. The sigmoidal function is reached by using an exponential equation and it is applied to normalize the sum of data inputs after being weighted.

ANN models

ANN consists of sets of input layer, hidden layer, output layer and each layer presented number of neurons connected with neurons in the adjacent layers through unidirectional connections. The information flow is only allowed in one direction during the training process, from the input layer to the output layer trough the hidden layer. The hidden layer has a synaptic weighting matrix and the weights are associated with all the connections made from the input layer to the hidden layer. There are two different arrangements of input and output of neurons:

- **Single Layer Perceptron (SLP)**: the arrangement of one input layer of neurons is fed forward to one output of neurons. It comprises of a single layer of weights and the inputs are directly connected to the outputs via a series of weights. The synaptic links connect every input to every output but only in a single way;

- **Multilayer Perceptron (MLP)**: between the input and output layers there are one or more intermediary layers called hidden layers. The computational units of the hidden layer are known as hidden neurons. The MLP can solve more complicated problems than SLP.
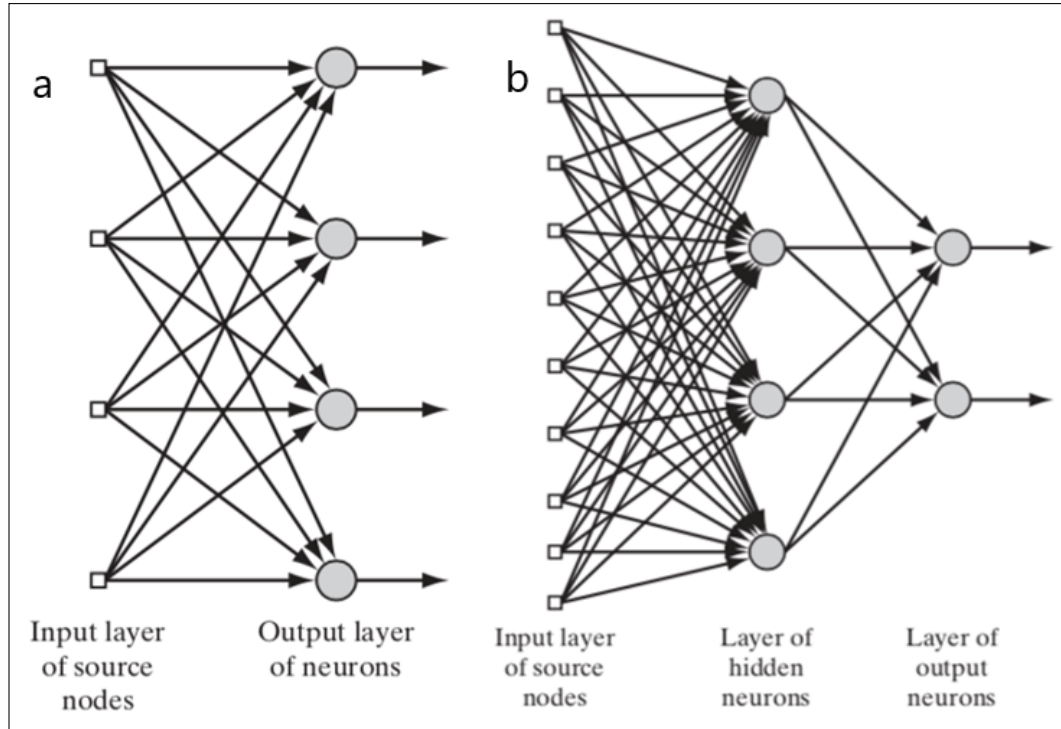
Figure 2.4: a: An example of single layer perceptron, b: an example of multilayer perceptron.

Backpropagation algorithm

Backpropagation algorithm [136] is the most popular of the MLP learning algorithms. Get h the hidden node, x the input node, w the weight and y the output node:

$$h_i = f \sum x_i w_{ij} \tag{2.12}$$

$$y_i = f \sum h_i w_{ij} \tag{2.13}$$

where i is the starting unit's identifier and j is the outcome unit's identifier. The error is the difference of the expected value t and the actual value k, and compute the error information term for both output and hidden nodes.

$$\delta y_i = y_i(1 - y_i)(t - y_i) \tag{2.14}$$

$$\delta h_i = h_i(1 - h_i)\delta y_i w_{jk} \tag{2.15}$$

δ_i the information error of the nodes. Finally, back propagate this error through the network by adjusting all of the weights, starting from the weights to the output layer and ending at the weights to the input layer.

$$\Delta w_{jk} = \eta \delta y_i h_i \qquad (2.16)$$

$$\Delta w_{ij} = \eta \delta h_i x_i \qquad (2.17)$$

$$w_{new} = \Delta w + w_{old} \qquad (2.18)$$

Where η is the learning rate. Backpropagation adjusts the weights in an amount proportional to the error for the given unit multiplied by the weight and its input. The training process continues until some termination criterion, such as a predefined mean square error or a maximum number of epochs.

Training algorithm for ANN

The training for ANN is not only based on memorizing the mapping relationship between inputs and outputs among the learning samples, but on extracting the internal rules about the environment which are hidden in the sample by learning the finite sample data. In my research activity, the MLP model which adopts the BP algorithm is developed. The forward-propagation input information is transferred to the output layer from the input layer after being processed in the hidden layer. The state of each layer neuron influences only the state of neurons in the next layer. If it does not obtain the expected output in the output layer, it shifts to back-propagation, and error signals are shown along the original pathway of the neural connection, in return the connection weight of each layer is modified one by one. Through successive iterations, the error between the expected output signals of the network and practical output signals of the system reaches an allowable range.

Advantages and disadvantages

The ANN has an excellent aptitude for learning the relationship between the input/output mapping from a given dates without any prior information or assumptions about the statistical distribution of data. This advantage makes the ANNs suitable for classification and prediction tasks in clinical situations. Moreover, the ANNs are inherently non-linear which makes them more useful for an accurate model of complex data patterns. In addition, the ANN has the ability to detect all the possible interactions between predictor variables. On the other hand, a disadvantage is that ANN presents difficulty to analyz systems which have many inputs due to a significant amount of time taken to the system. Another disadvantage is that ANN is prone to overfitting. This could be prevented limiting the number of hidden nodes, but it also reduces the power of the network to model complex non-linear relationships [137].

2.4.2   Classification and Regression Tree

Classification and Regression Tree Analysis (CART) [111] is a machine learning technique which is suited to the generation of clinical decision rules. It is a binary recursive partitioning technique: the term binary implies that group of elements, represented by a node, can only be split into two groups, the term recursive refers to the fact that the binary partitioning process can be applied over and over again and finally, the term partitioning refers to the fact that the dataset is split into sections or partitioned. Moreover, this algorithm is inherently non parametric that it means that no assumption is made regarding the underlying distribution of values of the predictor variables. As a classification problem, the CART consists of four main components. The first is the dependent variable that the algorithm hopes to predict, the second component is the predictors that are related to the outcomes variable of interest. The third component is the training dataset that includes values for both the outcome and predictor variables from whom the algorithm is able to predict outcomes. The last component is the test dataset from whom the algorithm is able to make accurate predictions.

Structure of CART algorithm

The CART algorithm consists of four steps:

- **Tree building**: it begins at the root node in which all data, of the learning dataset, are included. The algorithm finds the best possible variable to split the node into two child nodes checking all possible splitting variables as well as all possible values of the variables to be used to split the node. When the best splitter is chosen, the algorithm seeks to minimize the average purity of the two child nodes choosing, as the measure of purity, the splitting functions. The most common splitting function is the Gini index criterion which is computed as:

$$Gini\,index(t) = 1 - \left(\frac{n_i}{n}\right)^2 - \left(\frac{n_j}{n}\right)^2 \qquad (2.19)$$

  where t is the considered node, i and j are the two class labels, n_i and n_j are the number of subjects present at the node belonging either to the class i or j, respectively, and n is the total number of subjects at the node. The process of node splitting followed by the assignment of a predicted class to each node is repeated for each node and it continued recursively;

- **The stop tree building**: the process is stopped if there is only one observation in each of the child nodes, or if all observations within each child node have the identical distribution of predictive variables, making the splitting impossible, or if an external limit on the number of levels

in the maximal tree has been set by the users. The maximal tree which is created is generally over fit;

- **The tree pruning**: during this step, a sequence of simpler and simpler trees are generated, each of which is a candidate for the appropriately final tree that it is obtained applying the method of 'cost-complexity'. It is based on assigning a complexity penalty to the growth of large trees. Defining the cost complexity of tree $R_\alpha(\mathrm{T})$ as:

$$R_\alpha(T) = R(T) + \alpha|T|$$

  Where α is a real-valued scalar, $R(T)$ is a linear combination of misclassification probability and $|T|$ is the number of node. This method relies on a complexity parameter α which is gradually increased during the pruning process. Beginning at the last level, the child nodes are pruned away if the change in tree complexity is less than α times. So α is a measure of how much additional accuracy a split must add to the entire tree to warrant the additional complexity. As α is increased, more and more nodes are pruned away, resulting in simpler and simpler nodes;

- **The optimal tree selection**:the maximal tree fits the learning dataset with higher accuracy than any other tree but overestimating the performance. The aim is to select the optimal tree, defined with respect to expected performance on an independent set of data, and to find the correct complexity parameter α so that the information in the learning dataset is fit but not over fit. This means that the maximal tree will always give the best fit to the learning dataset.

Advantages and disadvantages

The advantages to apply this method are several, in particular this machine learning technique is very intuitive and easy to understand. Moreover, this algorithm needs less effort for data preparation during pre processing because it does not require either normalization or scaling of data. Another advantage is that missing values do not affect the process of building a decision tree. On the other hand, this structure presents some disadvantages such as it could be unstable, a small changes in the training dataset can cause a large change in the structure of the decision tree and it involves higher time to train the model and the training is relatively expensive as the complexity and time taken are more [138].

2.5   Feature selection

To select the inputs of the classification algorithms, several feature selection techniques were proposed in order to reduce the parameter from the original. The first technique selected the parameters correlated for less than 90%,

the second selection utilized the stepwise regression and the last one applied
the Principal Component Analysis (PCA). The stepwise regression is a semi-
automated process of building a model in which successively variables are
added or removed based on the t-statistics of their estimated coefficients. Given
a set of potential independent variables, the process extracts the best subset to
use in the forecasting model. The regression could start or with no variables
in the model and proceed adding one variable at a time, or with all poten-
tial variables in the model and proceed removing one variable at a time. At
each step, the program performs the following calculations: for each variable
currently in the model, it computes the t-statistic for its estimated coefficient,
squares it, and reports this as its "F-to-remove" statistic; for each variable not
in the model, it computes the t-statistic that its coefficient would have if it
was the next variable added, squares it, and reports this as its "F-to-enter"
statistic. At the next step, the program automatically enters the variable with
the highest F-to-enter statistic, or removes the variable with the lowest F-to-
remove statistic, the same significance level for both entry and exit tests is
recommended [139]. On the other hand, PCA, is a linear dimension reduction
tool that can be used to reduce large dataset of variables to a small set that
presented most of the information in the large set, minimizing the correlated
variables into a smaller number of uncorrelated variables called principal com-
ponents. These are new variables that are linear functions of those in the
original dataset that maximize variance and that are uncorrelated with each
other. These components are orthogonal, ordered such that the retention of
variation present in the original variables decreases as the order decreases, it
means that the first principal component retains maximum variation that was
present in the original components [140]. In addition to these techniques of fea-
tures selections, also all the parameters were considered as inputs of machine
learning techniques. Finally, the classification performances obtained adding
or not a specific clinical parameter, namely LVEF, were evaluated.

2.6 Classification performances

2.6.1 Confusion matrix

Starting from a basic confusion matrix of a binomial classification where there two classes (say, Y or N), the accuracy of classification of a specific example can be viewed one of four possible ways:

- True Positive (TP): the predicted class is Y, and the actual class is also Y;

- False Positive (FP): the predicted class is Y, and the actual class is N;

- False Negative (FN): the predicted class is N, and the actual class is Y;

- True Negative (TN): the predicted class is N, and the actual class is also N.

A basic confusion matrix is traditionally arranged as a 2x2 matrix as shown in Table 2.2. The predicted classes are arranged horizontally in rows and the actual classes are arranged vertically in columns, although sometimes this order is reversed [141].

		Observations	
		Y	N
Predicted Class	Y	True Positive Correct result	False Positive Unexpected result
	N	False Negative Missing Result	True Negative Correct absence of result

Table 2.2: Confusion matrix.

The statistical measures of the performance of a binary classification test are:

- **Sensitivity**: is defined as the ability of a classifier to select all the cases that need to be selected. Sensitivity (SEN)=TP/(TP+FN);

- **Specificity**: is defined as the ability of a classifier to reject all the cases that need to be rejected. Specificity (SPE)=TN/(TN+FP);

- **Precision**: is defined as the proportion of cases found that were actually relevant. Precision (PRE)=TP/(TP+FP);

- **Accuracy**: is defined as the ability of the classifier to select all cases that need to be selected and reject all cases that need to be rejected. Accuracy (ACC)=(TP+TN)/(TP+FP+TN+FN).

2.6.2   Receiver operator characteristic (ROC) curves and area under the curve (AUC)

Receiver Operator Characteristic (ROC) [142] curve is used to visualize the classifier's performance. It is created by plotting the fraction of true positives versus the fraction of false positives. When a table of such values is generated, the FP rate on the horizontal axis and the TP rate (same as sensitivity or recall) on the vertical axis were plotted. The FP can also be expressed as (1 − specificity). The Area Under the Curve (AUC) provides an aggregate measure of performance across all possible classification thresholds measuring the ability of the test to correctly classify the true positive and false positive. The AUC measures the entire two-dimensional area underneath the ROC curve. It takes values form 0 to 1.0, for the ideal classifier this quantity is 1.0, it means that the classifier is able to perfectly distinguish between the two classes. Instead a values of 0 indicates a perfect inaccurate classifier. An AUC of 0.5 suggests that the classifier has no discriminatory ability.

Chapter 3

Results and Discussions

This chapter, structured in three main sections, reported the main findings of my research activities. All the topics have the application of biomedical signal processing in common, providing useful information that clinicians can use to make decisions.

In the first section, the analysis of the circadian cardiovascular signals, such as blood pressure and heart rate, in normotensive and hypertensive subjects was conducted showing the same basic circadian rhythm with characteristic periods of time over 24hrs. Furthermore, since the relationship over 24hrs between these two signals could be useful to understand the control mechanism of the autonomic nervous system, the circadian BP/HR relation was studied applying linear regression analysis in the four characteristic periods. However, the difference between the BP and HR circadian rhythms affects their relationship, during each hour of the day. Since the conventional measurements (office and the average over 24hrs) did not accurately described the circadian changes due to the sympathetic/parasympathetic control system, in my research activity the linear approximation of BP/HR measurement in each hour of the day was examined and its variation along the 24hrs considered.

In the second section, the influence of age, gender and other risk factors (such as smoking, obesity and dyslipidemia) on BP and HR circadian rhythms has been discussed. At first, the effects of smoking on BP and HR signals in hypertensive and normotensive subjects, considering a temporal resolution of 15 minutes during day or 30 minutes during night, have been examined. Since the signals presented characteristic linear behaviors in different period of time over 24hrs, linear approximation in these intervals was evaluated associating their slopes to different risk levels of cardiovascular diseases as well as to the HR surge and decline over 24hrs. Concerning BP, it is known, from the literature, that this biomedical signal is influenced by the presence of cardiovascular risk factors as well as by ageing. Since, from a clinical point of view it can be relevant to understand and quantify how the relationship BP/Age is affected by risk factors in people with hypertension or not, linear regression analysis was applied considering values taken from office and ABPM measurements.

Finally, the influence of gender and age on the circadian rhythm of these cardiovascular signals, in subjects without hypertension, has been reported. In particular, the effect of age on the variation of heart beats over 24hrs was estimated in subjects divided in three age groups and how the gender influenced the circadian variation of BP/HR was analyzed.

In the last section, the research activity was focused on the application of parameters extracted from the variation of heart beats over 24hrs to identify subjects that presented different cardiovascular pathologies, in particular ischemic heart and dilated cardiomyopathy diseases. Two different mathematical approaches based on machine learning techniques for DSS based on HRV parameters as well as on age and gender have been proposed. The two algorithms concerning Artificial Neural Network based on generalized back-propagation and Classification and Regression Tree were developed. At first, DSSs were aimed at recognizing one single pathology, in particular, ANN being used to classify IHD and CART to identify DCM. Successively, CART was used to distinguish both IHD and DCM from normal subjects and between them, on a large dataset. The classification performances were, at first, evaluated using all the HRV parameters as input, then the algorithm was applied on features selected by using either stepwise regression, principal component analysis or correlation coefficients. Finally, in collaboration with clinicians and with the aim of only non-invasive parameters being considered, one clinical outcome (LVEF) was added to inputs to improve the classification performances.

3.1 Blood Pressure and Heart Rate circadian rhythms

In the first section, the results of two studies concerning the relationship between BP and HR signals were presented. First at all the circadian rhythm was examined in a large cohort of hypertensive and normotensive subjects. Successively, matching the subjects for age, the normotensive subjects were divided in normal/high normal and optimal BP subjects. Considering BPs and HR values each 15 min during the day and each 30 min during the night, the circadian trends were analyzed and four different time periods observed. In order to evaluate the relationship of BP and HR over 24hrs and to quantify the differences in circadian changes among subject groups, linear regression analysis was applied.

In the second section, since the single measure obtained in office or as the average along 24hrs did not accurately describe the circadian BP/HR changes, this relation was calculated hour-to-hour and the results were compared with those obtained considering BP and HR values in the conventional measurements.

3.1.1   Relationship between Blood Pressure and Heart Rate circadian rhythms and their differences among hypertensive and normal blood pressure subjects [1SG] [2SG]

Materials and methods

Subjects

Firstly, the cohort of subjects was composed of 629 subjects (261 males and 367 females, mean age 68.0 ± 16 years) in particular 423 Normotensive (NH) and 205 Hypertensive (H). Successively, matching by age, the dataset was reduced considering only 385 subjects (153 males, 232 females, mean age 65.0 ± 15.6 years) divided as 215 Hypertensive (H), 112 Normal/high normal blood pressure (NHn) and 58 Optimal Blood pressure (NHob) subjects.

Statistical analysis

The circadian behavior of the mean values of SBP, DBP and Mean SBP (i.e. [2*DBP+SBP]/3) among the subjects was separately examined for NH and H subjects. The HR circadian behavior was compared in the two groups using the Wilcoxon signed rank sum test, assumed that the subject groups were independent and showed a non-Gaussian distribution. In the second study, to evaluate the significance of the differences between each pair of the three groups, the Wilcoxon rank sum test with the Bonferroni correction was applied. Since the BP profiles during 24hrs showed to be bimodal with two maxima and two minima (Figure 3.1), the 24hrs were divided in four intervals. In each period the quite linear trend was fitted by a regression line and the linear approximation goodness was measured by using the R-square statistic.

Results

The BPs presented a similar circadian trend in NH and H subjects and an about constant difference of 20mmHg along 24hrs. The circadian rhythm of HR (Figure 3.2) showed slow decrease during the day from about 9:30 until 19:30 and a quicker decrease from 19:30 till 2:00. From 2:00 to 6:00 the HR was about constant followed by a quick increase during the morning from about 6:00 till 9:30. The HR values were significantly higher (p<0.001) in H than in NH during all 24hrs with a difference between H and NH of about 2.3±0.8 bpm. The relationships between SBP, Mean BP, DBP and HR circadian rhythms in NH and H subjects showed that between 9:30 to 15:00 (blue line), the HR and SBP decrease progressively while during postprandial hours, between 15:00 to 19:30 (red line), SBP increases and HR still decreases. During the first part of the night (from 19:30 to 2:30), HR and SBP show a marked reduction with the presence of hysteresis in respect of the quick increase of both SBP and HR values during the successive morning period, from 2:30 to 9:30 (green line).
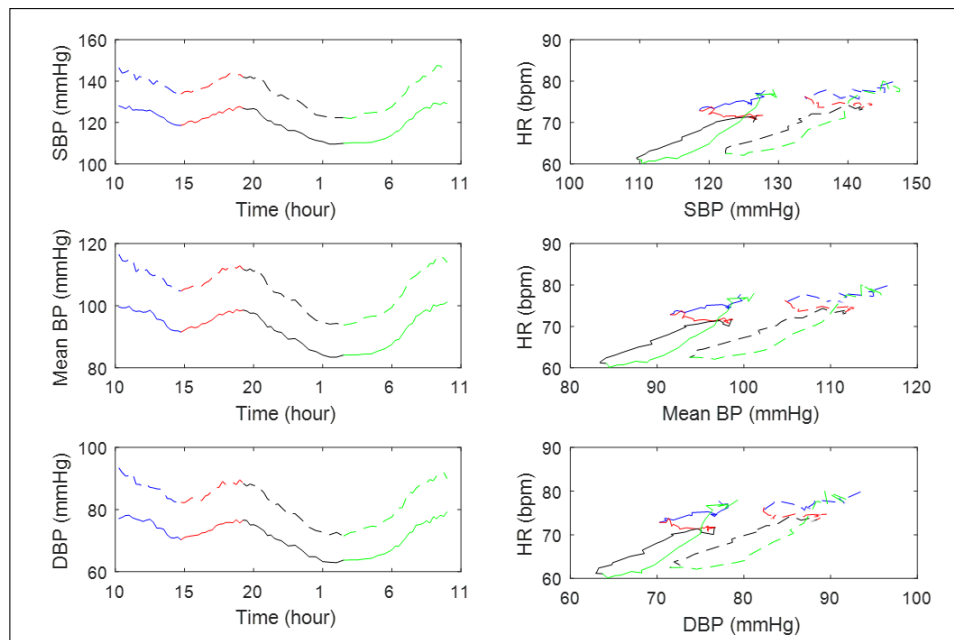
Figure 3.1: Mean circadian rhythms of SBP, Mean BP, DBP (left panels) and relationships between HR and SBP, Mean BP and DBP (right panels) in Normotensive (solid line) and Hypertensive subjects (dashed line).
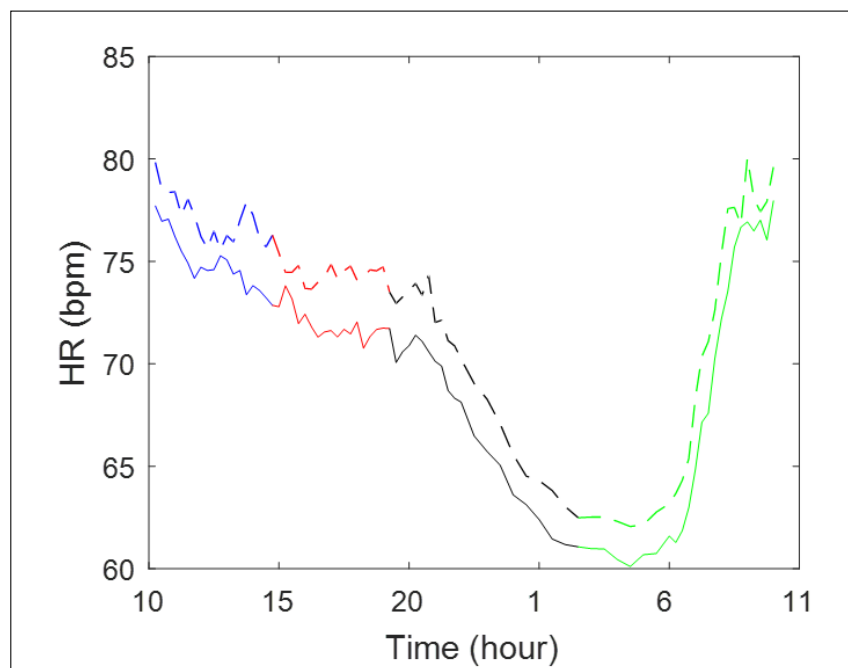
Figure 3.2: Mean circadian rhythm of HR in Normotensive (NH, solid line) and Hypertensive (H, dashed line) subjects.
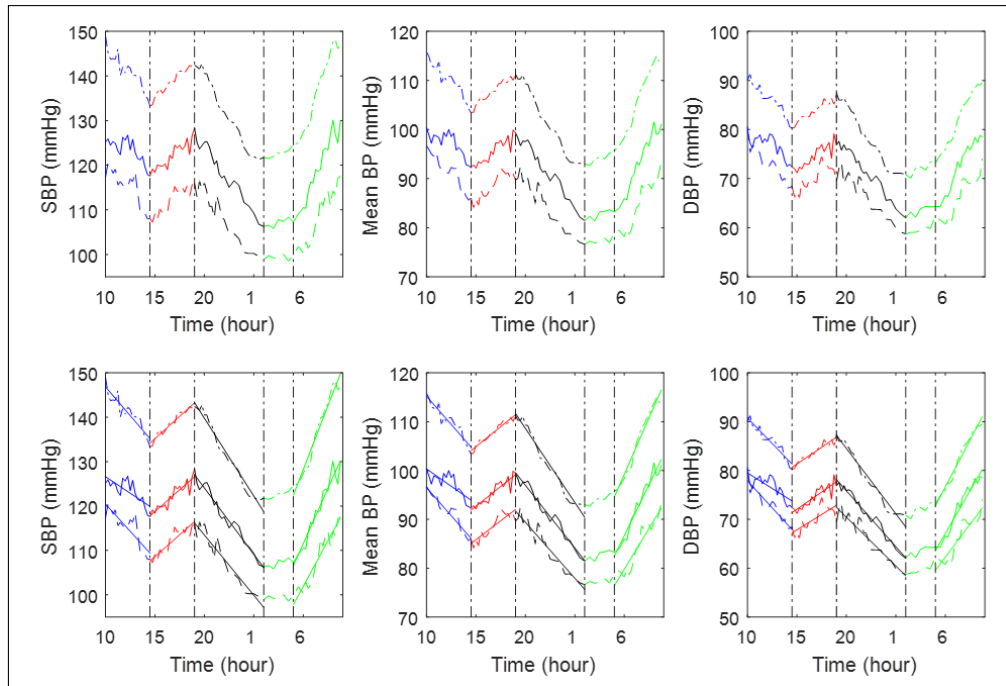
Figure 3.3: Top panels: mean circadian rhythms of SBP, Mean BP and DBP in H (dot-point line), NHn (solid line) and NHob (dashed line) subject groups, subdivided in four periods: 10:00-14:30; 14:30-19:00; 19:00-2:00; 5:00-10:00. Bottom panels: Circadian rhythms as in the top panels together with linear regression curves.

The linear relation (Figure 3.3), which approximated the trend of the BPs values in each of the four periods, showed high values of the R^2 (R^2 77%–98% in NHn, R^2 75%–96% in NHob and R^2 75%–99% in H), validating the linear approximations, and a significant relation (p-values < 0.0001) with the time. Between 10:00 and 14:30 the slopes of SBP, Mean BP and DBP, were higher in H group than NHn group. From 14:30 to 19:00, the BP slopes in H group were higher than in the others two groups and BP increased more slowly in NHob than in NHn. In the two decreasing and increasing tracts between 19:00 and 10:00, the slopes were higher in H than in NHob and NHn being very similar in the latter (Table 3.1). Figure 3.4 shows the circadian rhythm of HR in the three groups with no significant differences in the behavior between NHn and the other two groups. Considering the linear approximation between 19:00 and 2:00, the slopes of the three groups were similar and lower than from 5:00 to 10:00. The corresponding two linear approximations were validated by the high values of R^2 (R^2 95% in NHn, R^2 87% in NHob and R^2 93% in H).

		Slope				Intercept			
		10:00-14:30	14:30-19:00	19:00-2:00	5:00-10:00	10:00-14:30	14:30-19:00	19:00-2:00	5:00-10:00
SBP	H	-2.59	2.01	-3.50	5.74	172.78	104.90	209.96	44.13
	NHn	-1.49	1.96	-2.99	4.97	141.59	89.23	183.85	-37.81
	NHob	-2.45	1.97	-2.71	4.11	145.03	78.99	167.82	-21.47
Mean BP	H	-2.33	1.63	-3.03	4.61	138.37	80.37	169.40	-39.28
	NHn	-1.42	1.62	-2.57	4.16	114.57	68.36	148.45	-38.27
	NHob	-2.31	1.54	-2.29	3.34	119.67	62.66	135.33	-20.71
DBP	H	-2.09	1.36	2.68	3.84	111.74	60.86	138.20	-38.87
	NHn	-1.27	1.40	-2.29	3.50	92.27	51.02	121.65	-38.17
	NHob	-2.21	1.25	-1.99	2.69	100.00	48.97	110.41	-18.55
HR	H	-0.81	-0.51	-1.72	4.64	87.44	82.99	107.63	-75.26
	NHn	-0.51	-0.36	-2.03	5.14	83.39	80.57	114.48	-91.22
	NHob	-0.08	-0.38	-1.87	4.56	74.03	80.03	109.79	-76.14

Table 3.1: Slope (in mmHg/hour for BPs and bpm/hour for HR) and Intercept (in mmHg for BPs and in bpm for HR) values of the linear regression lines for SBP, Mean BP, DBP and HR.
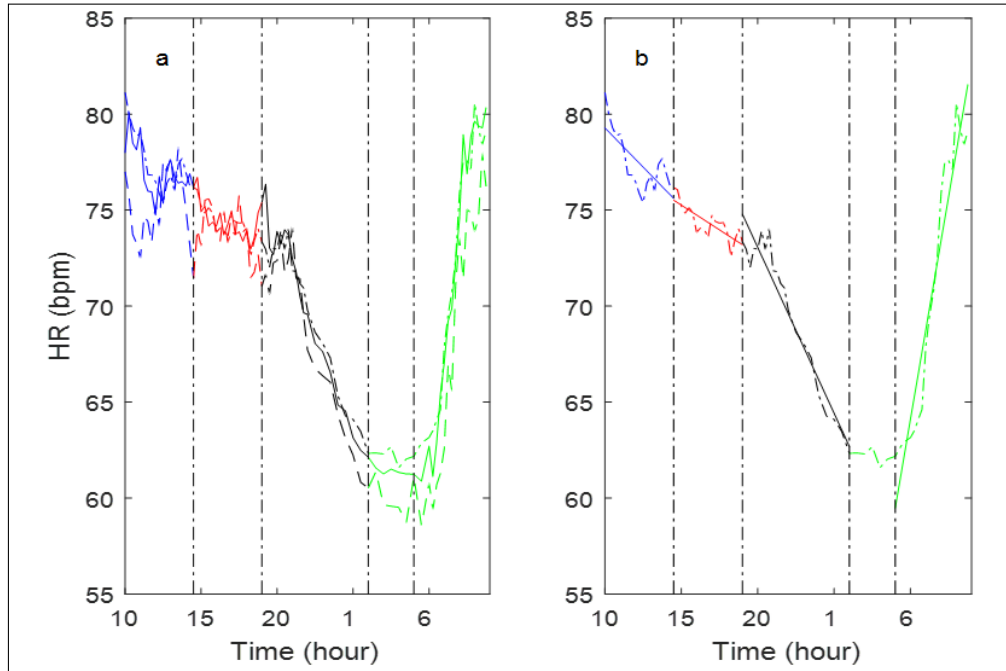

Figure 3.4: a: Mean circadian behaviour of HR values in H (dot point line), NHn (solid line) and NHob (dashed line) subjects. b: Mean circadian behaviour of HR values in H subjects together with the linear regression curves. The 24hrs were divided in four periods: 10:00-14:30; 14:30-19:00; 19:00-2:00; 5:00-10:00.

Discussion

The circadian behaviors of BP ad HR were similar both in H and NH subjects. The changes in HR present closely parallel changes in BPs with a marked reduction during nocturnal rest. The results confirmed previous outcomes concerning generally lower values of HR and BP during the night than during the day [16, 17] as well as a direct association between HR and SBP and DBP increasing [14] at least in period in which HR increased (2:30-9:30). In the interval between 15:30 and 19:30, HR and BP showed an inverse relationship with decreasing HR and increasing BP (Figure 3.1, left panels) not yet reported in the literature. Furthermore, a true linear relationship between HR and BP values was found only during the night with a decrease of BP (from 19:30 to 2:30), partially confirming previous literature results [14]. In the other three periods of the day, the relationship changed drastically and was mostly not linear. Studying accurately the NH subjects, the circadian rhythms of HR and BP presented specific trends with lower values during the night than the day [5, 10, 15, 16, 17]. Moreover, the circadian BP and HR behaviors in H presented the same profile as in NHn and NHob subjects partially confirming the results of [27] only in H subjects.

Main finding

In these studies the highest values of slopes were presented during awakening, particularly in the H group while the slowest values were found in the post prandial period suggesting that the mechanisms that regulate the circadian behavior of the two signals are independent. Applying linear regression technique, the slopes could be used to quantify the rate of change of the morning blood pressure surge, associated with acute cardiovascular events.

3.1.2 Influence of the time of day on the relationship between Heart Rate and Blood Pressure [3SG]

Materials and Methods

Subjects

The study population included 388 subjects (155 males and 233 females, aged 65±16 years old) composed of 216 Hyperthensive (H) and 172 Non Hyperthensive (NH) subjects.

Stastical analysis

The BP and HR values were averaged hour-to-hour in each subject. The slope, the intercept and the R^2 of a regression line, fitting the relation between BPs and HR data, considering all the subjects divided in the two groups (NH and H), were evaluated for each hour. The relationship between BP and HR was also evaluated considering, for each subject, either the office measurements or the average along the 24hrs. Also in this case, the regression lines fitting the

SBP/HR and DBP/HR relationships were separately estimated for NH and H subject groups. The significance of the presence of the linear relation was tested statistically comparing each slope with the zero value.

Results

The slope values of the hour-to-hour behavior of SBP/HR (Figure 3.5) were significantly ($p<0.05$) different from zero for almost all the hours in NH subjects while in H patients, from 01:00pm to 02:00pm and from 10:00pm to 11:00pm and at 07:00am the p-values presented significant values (Figure 3.6).
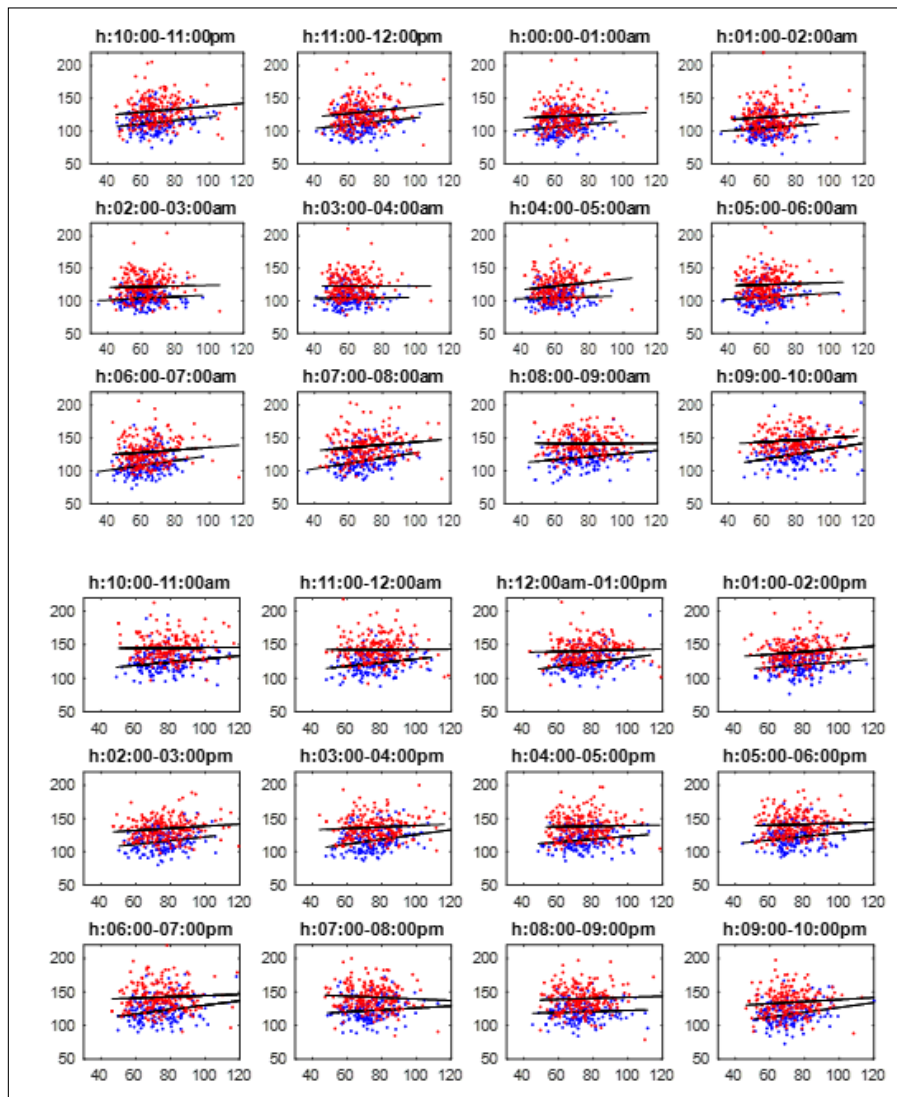

Figure 3.5: Relationships between SBP (vertical axis, in mmHg) and HR (horizontal axis, in bpm) for every hour along the 24hrs, from 10:00am to 09:00am. Red points: hypertensive patients; blue points: normotensive subjects; black lines: regression lines of the SBP/HR relationships in the two subject groups.
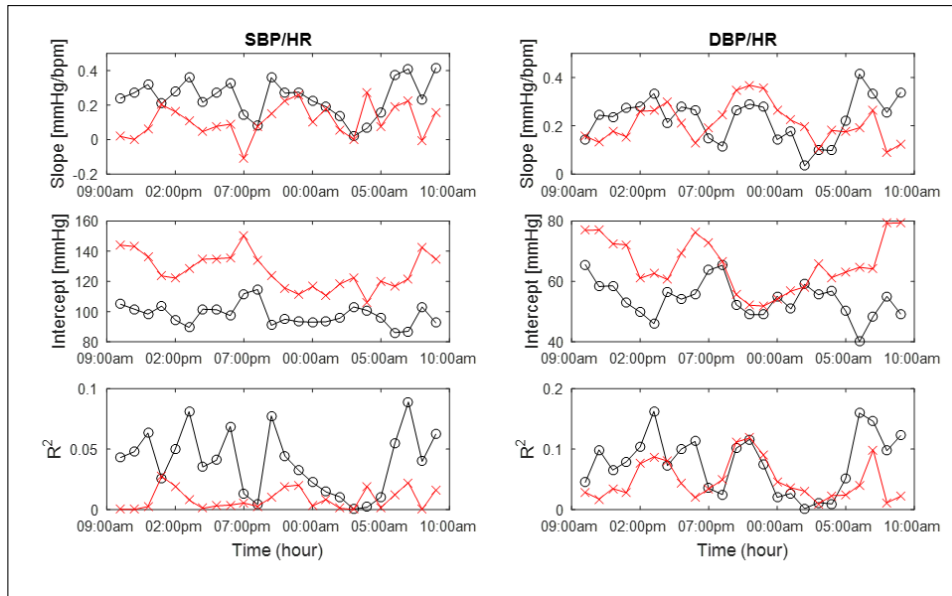
Figure 3.6: Slope, intercept and R-square values of the hour-to-hour regression lines of the BP/HR relationships (left panels: SBP/HR, right panels: DBP/HR) in the two subject groups during 24hrs. Red crosses: hypertensive patients; black circles: normotensive subjects.

The hour-to-hour relationship between DBP and HR had corresponding slope values significantly different from zero ($p<0.05$) in almost all the hour except from 02:00am to 04:00am in NH and at 03:00am and at 08:00am in H subjects (Figure 3.6). From 05:00am to 07:00pm, the slope presented significantly ($p<0.0001$) higher values in NH than in H subjects for SBP/HR relationship with a similar behavior also for DBP/HR relation. During the night, the slope of SBP/HR presented comparable values in both subject groups while DBP/HR presented significant ($p<0.0001$) higher values in H. The difference between the two groups was significant during the day in both relationships while during the night the difference was reduced for SBP/HR and almost nulled for DBP/HR relation. For SBP/HR, R^2 was always lower than 0.02 in H group, highlighting a very large inter-subject variability while in NH group, the variability was high during the night. The R^2 for DBP/HR were always greater than 0.01 for H and NH during the day, instead comparable values between subject groups were presented during the night. Figure 3.7 shows that, for SBP, the slopes in H were negative while positive values were found in NH (Table 3.2); under no circumstances the slopes were significantly different from zero (no relation between HR and SBP) although the difference between the two groups was significant ($p<0.0001$) in both cases. Considering the DBP, the slopes in both subject groups were positive (Table 3.2), always significantly different from zero ($p<0.01$) except for H in office condition. The difference between slopes in the two subject groups was significant ($p<0.01$) in both measurements. Moreover, the intercept values, both for SBP/HR and

DBP/HR, were greater in office than in the average on 24hrs in both subject groups (Table 3.2) and significantly (p<0.001) different between them except in the average on 24hrs for DBP/HR.

		Office		24hrs	
		H	NH	H	NH
SBP	Slope [mmHg/bpm]	-0.09	0.07	-0.1	0.1
	Intercept [mmHg]	157	114	145	107
	R^2	0.0001	0.006	0.008	0.01
DBP	Slope [mmHg/bpm]	0.07	0.1	0.2	0.2
	Intercept [mmHg]	90	67	69	57
	R^2	0.008	0.04	0.03	0.06

Table 3.2: Slope, intercept and R^2 values of the linear approximations of Fig.3.7.
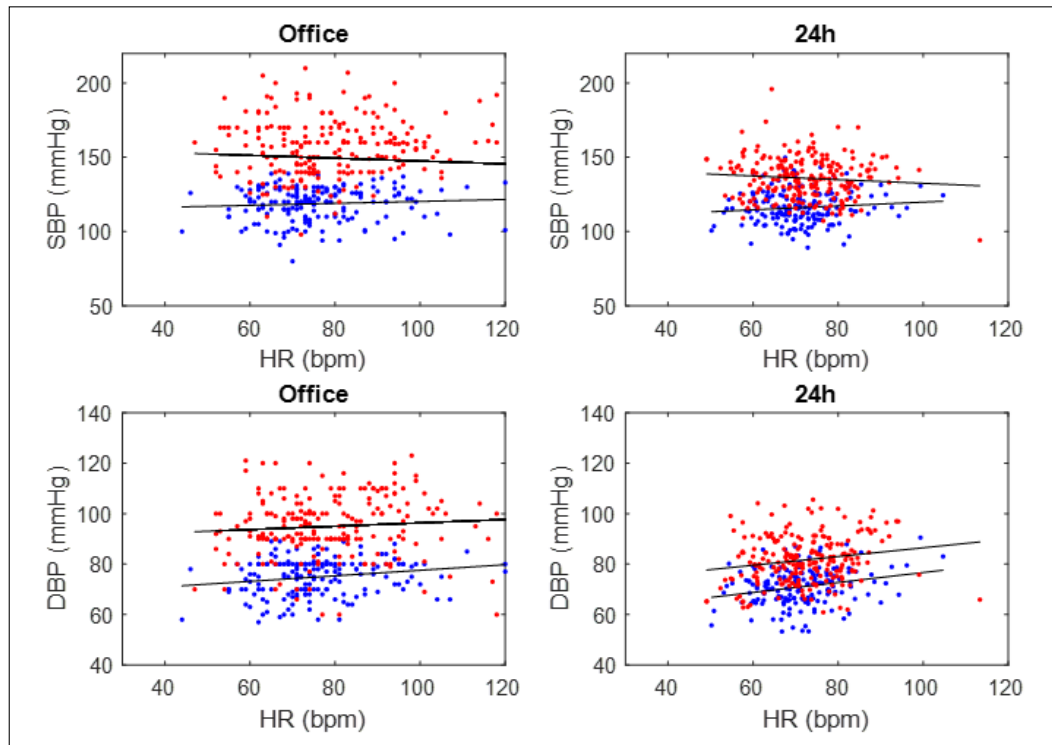

Figure 3.7: SBP/HR and DBP/HR relationships in the two subject groups (red points: hypertensive patients; blue points: normotensive subjects) calculated in office condition (left panels) and as average on 24hrs (right panels).
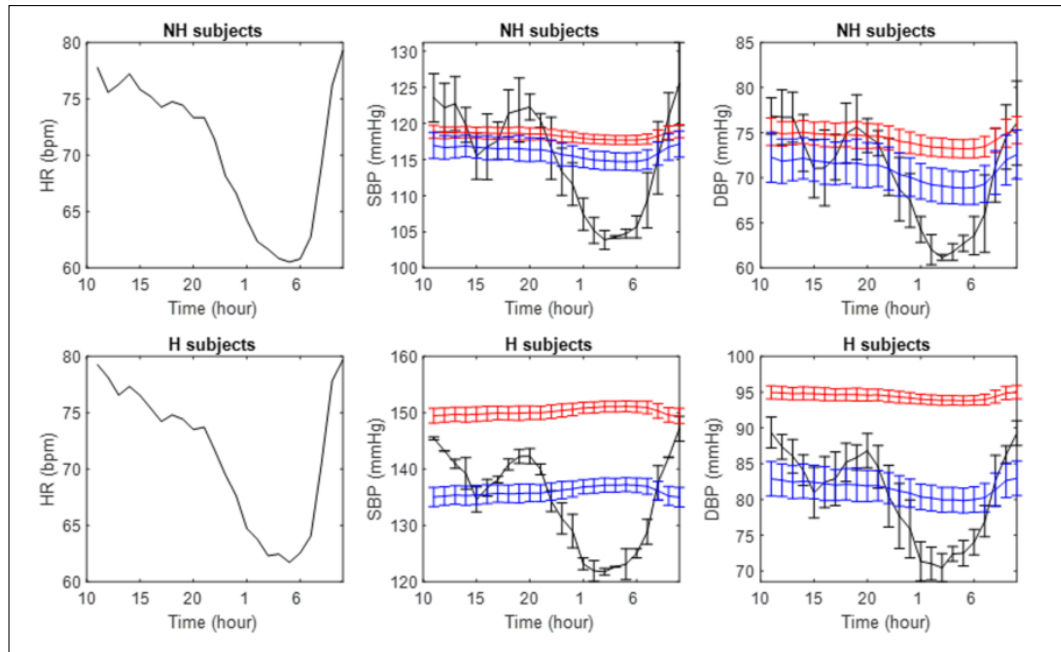
Figure 3.8: Left panels: HR profiles in the two subject groups (NH and H); Central and right panels: SBP and DBP circadian profiles estimated from HR profile by using slope and intercept values calculated either in the hour-to-hour analysis of Figure 3.7 (black line), or in the office (red line) or in the average on 24hrs (blue line) of Table 3.2. The bars correspond to ± 1 Standard Deviation among the subjects.

Finally, the R^2 for SBP was in both groups always lower than 0.01 while for DBP it was always lower than 0.06. To compare the reliability of the linear approximation estimated in the different ways such as hour-to-hour, office and average on 24hrs, the circadian BPs profiles were estimated form HR using the slope and intercept evaluated in Figure 3.8. The rhythms evaluated starting from office as well as from averages on 24hrs relationships, were very poor, especially using office values in H subjects (Figure 3.8). Furthermore, the circadian variability along 24hrs in these two situations was markedly reduced in comparison of the correct circadian rhythm; in H the trend was even opposite. In addition, as expected since the slope and intercept are unique for all the hours, the circadian profiles were similar to that of the HR rhythm rather than to that of SBP or DBP.

Discussion

The slopes for SBP/HR and DBP/HR were significantly different from zero almost exclusively in NH showing the greatest values during the morning. This fact supports the hypothesis of a negligible relation between HR and both SBP and DBP and that the two systems controlling cardiac and pressure variability could work independently in night time period as well as during most of the day for H. The variability (measured as the standard deviation) of

slopes and intercepts along the 24hrs were quite large and similar in the two groups and for SBP/HR and DBP/HR relations (about 0.08-0.11mmHg/bpm for slopes and 6-12mmHg for intercepts). This may suggest that only one punctual measure (as that in office or as the average on 24hrs) could be not adequate to describe the effective link between BP and HR during the day in each category of subjects. The circadian rhythms of SBP and DBP in H and NH, estimated using the linear regression obtained starting from office as well as from average on 24hrs relationships, very poorly approximate the correct profile (Figure 3.8), especially using office values in hypertensive subjects. By using the hour-to-hour slope and intercept estimations is possible to obtain, from HR values, good approximation values for both SBP and DBP for all 24hrs. In fact, mean values of SD bars in Figure 3.8 were of about 2.9mmHg in NH and 1.5mmHg in H subjects for SBP and of 2.7mmHg in NH and 2.6mmHg in H subjects for DBP. This fact underlines the limitation due to the use of a single measure (office or 24hrs average) to explain the relationship between the two variables that changes during 24hrs in a different way in the two populations and that HR is not the only determinant of SBP and DBP changes in the circadian rhythm. In conclusion, for a correct estimation of both the relationships BP/HR and of BP values from HR ones along the 24hrs, it is suggested to use linear approximations extracted from the hour-to-hour analysis rather than those used so far.

Main findings

The circadian trend of the relationship between BPs and HR was evaluated applying a regression line for each hour along the 24hrs considering the two subject groups. The slopes presented values near zero in the night supporting the hypothesis that the two system controlling cardiac and pressure variability could work independently. Moreover, the large variability of slopes and intercepts along the 24hrs suggests that the punctual measurement either in office or as the average over 24hrs are not sufficient to describe the BPs/HR relationship.

3.2 Influence of age, gender and other risk factors on Blood Pressure and Heart Rate circadian rhythms

In this section, the effects of smoking on cardiovascular circadian rhythms in hypertensive and normotensive subjects, are reported. To evaluate the influence on HR, the linear approximation of the circadian behavior was applied in specific characteristic periods of time. Since no large differences were found in the trend between 10:00-14.30 and 14:30-19:00 in both subject groups, the 24hrs were divided in only three periods of time, from late morning to evening (10:00-20:00), from evening to night (20:00-04:00) and from early morning to

late morning (05:00-10:00). Furthermore, the effect of smoking on the circadian BP linear trend was evaluated, dividing the 24hrs in four periods, due to the bimodal behavior of BP. In that case, the subjects which presented other risk factors (such as obesity and dyslipidemia) were not considered and the cohort of subjects reduced. However, there are other risk factors such as obesity and dyslipidemia, that affected the circadian rhythms and the periods of time along the 24hrs, in which the signal's alterations were manifested, could be associated with different risk levels of cardiovascular diseases. Thus, the influence of each of these risk factors on HR circadian rhythm was examined in normotensive and hypertensive subjects applying a linear regression on the three periods over 24hrs. Successively, since age increases BP, the relationship between BP and age was examined considering the presence of other risk factors, evaluating and comparing the differences between office and ABPM measurements. Moreover, the influence of age on the circadian rhythm of HRV in subjects between 15 and 90 years old was evaluated and a possible model of this relation was estimated. Finally, how the relationship between BP and HR depends on gender, their relation over 24hrs was accurately described, using both office and ABPM measurements.

3.2.1 Influence of smoking on Heart Rate and Blood Pressure circadian rhythm in hypertensive and non-hypertensive subjects [4SG] [5SG]

Material and Methods

Subjects

To evaluate the effect of smoking on HR, the cohort was composed of: 58 Hypertensive Smoker (HS) (36 female and 26 male, 56±14 mean age), 351 Hypertensive Non Smoker (HNS) (223 female ad 128 male, 66±15 mean age), 39 Non Hypertensive Smokers (NHS) (26 female and 13 male, 53±15 mean age) and 166 Non Hypertensive Non Smokers (NHNS) (92 female and 74 male, 63±15 mean age). Since the presence of other several cardiovascular risk factors such as obesity, dyslipidemia and diabetic mellitus could affect the BP circadian rhythm, in the second study the subjects without those factors were selected. Hence, the cohort was composed of 32 Hypertensive Subjects (HS) (18 male and 14 female, 52±14 mean age), 113 Hypertensive Non Smokers (HNS) (58 male and 55 female, 56±13 mean age), 20 Non Hypertensive Smokers (NHS) (13 male and 7 female, 54±15 mean age) and 83 Non Hypertensive Non Smokers (NHNS) (46 male and 37 female, 59±16 mean age).

Statistical analysis

The 24hrs HR profiles showed three specific intervals corresponding to: late morning to evening (10:00-20:00), evening to night (20:00-04:00) and early morning to late morning (05:00-10:00). In each period the quite linear trend

was fitted by a regression line and the slopes and intercepts were evaluated. In the second study, the same technique was applied to study mean values of BPs for each subject groups. Since the BP profiles during 24hrs presented a bimodal trend, the 24hrs were divided into four intervals and in each period the quite linear trend was fitted by a regression line. Finally, the trends between normotensive smokers and non smokers and between hypertensive and non-smokers were compared in each interval and for both signals using the Wilcoxon rank sum test, assuming that the subject groups were independent and showed a non-Gaussian distribution.

Results

All the linear HR trends decrease slowly between 10:15 and 19:45 (day time), then decrease quickly from 20:00 to 4:00 (night time) and finally they increase very quickly during the early morning (5:00-10:00) (Figure 3.9, Table 3.3).
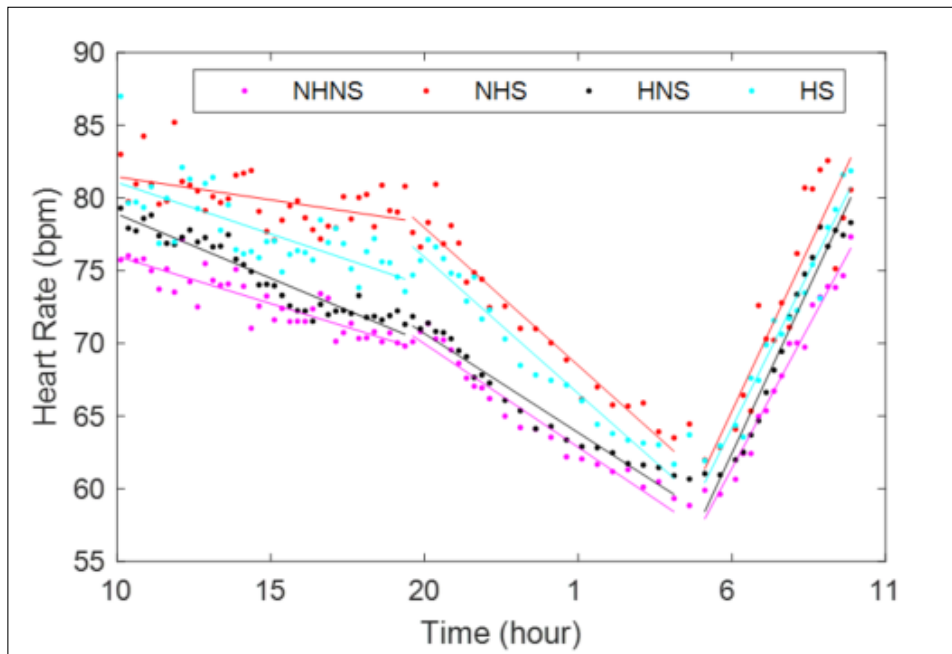
Figure 3.9: Mean circadian HR rhythms in Hypertensive Smokers (HS), Hypertensive Non Smokers (HNS), Non Hypertensive Smokers (NHS) and Non Hypertensive Non Smokers (NHNS) together with the linear approximations in the three separate periods.

The differences between the HR circadian trends in smokers and non-smokers were significant and about constant during the 24hrs both in H (mean value 2.8±2.0 bpm, p-value<10-10) and in NH (mean value 6.4±2.3 bpm, p-value<10-10) subjects. On the other hand, the differences between circadian SBP and DBP values in HS compared with HNS were significant in the three periods between 10:00 and 02:00 (p<0.05)(Table 3.4). Moreover, in NH, the difference between smokers and non-smokers was significant only for SBP values

between 10:00 and 19:00 (p<0.04). The BPs behaviours were approximated, by linear regressions (Figure 3.10) presenting slopes that differ among groups and intervals (Table 3.5).

Groups	Age (years)	Gender		Periods during 24hrs								
				10:15-19:45			20:00-04:00			05:00-10:00		
		F	M	Mean±SD	Slope	r	Mean±SD	Slope	r	Mean±SD	Slope	r
NHS	63±15	92	74	72.9±2.1	-0.65	0.88	65.4±3.9	-1.42	0.98	68.1±5.6	3.92	0.99
NHS	53±15	26	13	80.0±1.8	-0.32	0.49	71.9±5.2	-1.89	0.97	73.0±6.9	4.51	0.91
HNS	66±15	223	128	74.7±2.6	-0.88	0.95	66.3±3.8	-1.36	0.98	70.2±6.6	4.55	0.98
HS	56±15	36	26	77.7±2.7	-0.71	0.74	69.9±5.1	-1.87	0.98	71.5±6.2	4.28	0.98

Table 3.3: HR mean values (±SD) (bpm), slopes (bpm/hour) and correlation coefficient of the linear approximation in each of the three periods of the circadian profile for the four groups.

Systolic BP	NHNS	NHS	HNS	HS
10:00-14:30	124-128#	121-126#	137-143*	142-146*
14:45-19:00	124-126#	121-125#	134-137*	136-144*
19:15-02:00	116-127	112-125	121-138*	126-144*
02:30-09:45	113-122	109-122	120-139	120-134
Day time	122-127§	120-125§	134-140$	136-145$
Night time	113-116§	107-113§	118-121	118-125
Dyastolic BP	**NHNS**	**NHS**	**HNS**	**HS**
10:00-14:30	77-81	77-81	88-90*	88-93*
14:45-19:00	76-78	77-81	83-86*	85-91*
19:15-02:00	68-79	71-79	72-85*	75-91*
02:30-09:45	66-75	68-79	73-86	72-84
Day time	75-80	77-81	83-88$	85-92$
Night time	66-68	67-71	69-74	70-77

Table 3.4: 25th and 75th percentiles of BP pressures (in mmHg) in the four subject groups for each period and during day- and night times. Differences between smokers and non-smokers: # p<0.04, * p<0.05, § p<0.003, $ p<0.0001.

	NHNS		NHS		HNS		HS	
Systolic BP	m	q	m	q	m	q	m	q
10:00-14:30	-2.01	151	-0.52*	130	-3.13	179	-2.32	173
14:45-19:00	1.03	108	2.04	88	1.06	117	2.92	90
19:15-02:00	-2.62	180	-3.71	201	-3.77	213	-4.25	229
02:30-09:45	2.12	53	2.77	31	4.32	-4	3.71	14
Dyastolic BP	m	q	m	q	m	q	m	q
10:00-14:30	-1.67	99	-0.58*	86	-2.37	118	-2.15	118
14:45-19:00	1.13	59	1.69	50	1.13	65	2.95	38
19:15-02:00	-2.52	129	-2.34	127	-3.28	152	-4.04	173
02:30-09:45	2.15	6	2.24	6	3.39	-24	3.22	-20

Table 3.5: Slopes (m in mmHg/hour) and intercepts (q in mmHg) of the regression lines between hours of the day and BP, in the four groups for each period. *No significantly different from zero.
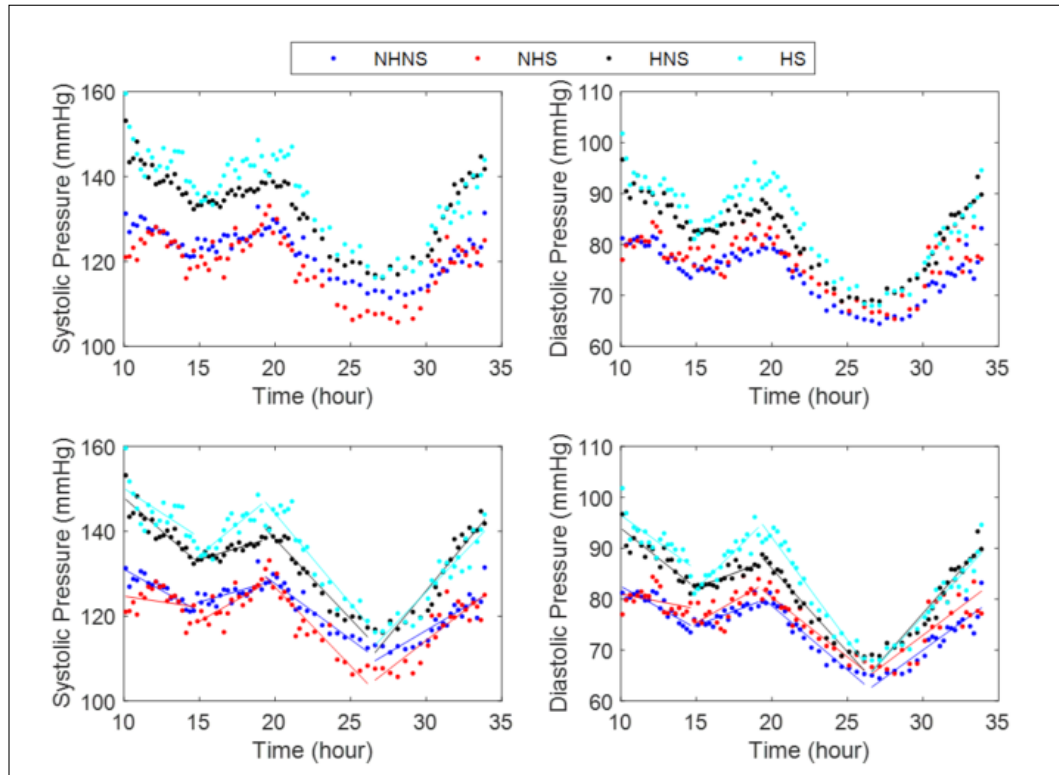

Figure 3.10: Circadian Systolic and Diastolic BP rhythms in Hypertensive Smokers (HS), Hypertensive Non-Smokers (HNS), Normotensive Smokers (NHS) and Normotensive Non-Smokers (NHNS) together with the linear approximations in the four separate periods.

The slopes were generally greater for SBP in smokers than in non-smokers (in H and NH subjects) in the three periods after 14:45 while an opposite trend was present between 10:00 and 14:30 with slope values greater in non-smokers (Table 3.5). For DBP the slopes were quite similar among the subject groups, excluded HS, in the three periods after 14:45 while from 14:45 to 02:00 in the HS group, the slopes were much higher. For BPs between 10:00 and 14:45 the slopes were greater in non-smokers than in smokers. R^2 presented values between 0.40 and 0.95 for both SBP and DBP, confirming a good linear approximation. The only interval in which the BP presented a quite constant behaviour was between 10:00 and 14:30 in NHS. In all the other cases the slopes were significantly (p<0.01) different from zero

Discussion

In all the tree periods selected, smokers presented significant higher HR values than non-smokers with NHS subjects showing the highest values and NHNS subjects the lowest (Figure 3.9). The differences remained quite constant during the 24hrs with higher values in NH subjects supporting the hypothesis that smoking has a larger impact on this category of subjects confirming the outcomes reported by [36, 37]. The analysis using linear approximation highlights different velocity changes of HR during day time, in the second period similar slopes in HS and NHNS, and during night time greater changes velocity of HR. A similar behavior was presented by non-smokers but with slightly lower slopes. In the last period, a very similar trend among all subjects demonstrated a very quick increase of HR, not associated with either the smoking or the hypertension, probably due to the increase of metabolic and oxygen needs during the early morning. Considering BPs along 24hrs, normotensive subjects presented higher mean values of SBP in non-smokers than in smokers partially confirming [37]. HS presented significant higher values than HNS along the periods, between 10:00 and 02:00, while during the awaking the behavior was opposite. These results confirm the findings of numerous authors [35, 36] on day time, enlarging the time interval in which this occurs. The direct pharmacological effects of nicotine in different daily activities reduced during the night time could explain this fact. Using linear approximation it was observed that in the first period the slopes were different among the four subject groups, showing different velocity changes of BPs with higher values in both hypertensive and normotensive non-smokers than in smokers. In the successive three periods, the smokers had slopes generally greater than non-smokers in both H and NH subjects. The values were similar between smokers and non-smokers only during both the third period for DBP in NH and the fourth period, the awaking, for SBP in H. This behavior could underline that during awakening, BP variations were mostly regulated by the autonomic nervous system that minimizes the effect of smoking. There were significant differences between smokers and non smokers in hypertensive and non hypertensive during the day time with higher rates in non smokers between 9:00 and 14:30 and in smokers in the other three periods. In conclusion, the different linear velocity rates of HR and BP

changes during 24hrs could be linked to different risk levels of cardiovascular diseases.

Main findings

Smoking affected both HR and BP circadian rhythms over 24hrs. In particular higher linear velocity rates of HR were found during day and night time in smokers than in non smokers, both in normotensive and hypertensive subjects. Considering BP, between 10:00 and 14:30 the changing rate was higher in non-smokers than in smokers for both normotensive and hypertensive subjects while an opposite behavior was found from 14:30 to 10:00. The different velocity rates could have a prognostic meaning because they could be associated with different risk levels of cardiovascular diseases

3.2.2 Influence of smoking and other cardiovascular risk factors on Heart Rate circadian rhythm in normotensive and hypertensive subjects [6SG]

Materials and methods

Subjects

It was considered a sample of 618 subjects composed of: 83 Normotensive without risk factors NH (46 male and 37 female, 59±15 mean age), 20 Normotensive Smoker NHS (13 male and 7 female, 54±15 mean age), 44 Normotensive with dyslipidemia (NHD) (23 male and 21 female, 70±11 mean age), 23 Normotensive obese (NHO) (95 male and 74 female, 61±17 mean age), 169 Hypertensive without risk factors (H) (95 male and 74 female, 64±16 mean age), 32 Hypertensive Smoker (HS) (18 male and 14 female, 52±14 mean age), 99 Hyperthensive with dyslipidemia (HD) (73 male and 26 female, 70±13 mean age) and 53 Hyperthensive obese (HO) (34 male and 19 female, 65±14 mean age).The information about risk factors were collected in accordance with international guidelines [143].

Statistical analysis

The 24hrs HR profiles were subdivided in three intervals corresponding to: late morning to evening (10:00-20:00), evening to night (20:00-04:00) and early morning to late morning (05:00-10:00). In each period, a regression line fitted the linear trend and the linear approximation significance was measured by using the R^2 and the p-value. Moreover, the mean and standard deviation values of HR among the subjects was separately examined in each period for each groups. In each period, the differences between the HR mean values calculated in H and in each of other normotensive groups (NHS , NHD and NHO) as well as between H and each of other hypertensive groups were evaluated (HS, HD and HO). Furthermore, to assess the influence of hypertension on HR rhythm when the same risk factor was present. The differences were estimated between

each pair of normotensive and hypertensive corresponding groups (NH vs H , NHS vs HS, etc.). In order to calculate these pair comparisons, the Wilcoxon rank sum test, assuming that the subject groups were independent and showed a non-Gaussian distribution, with the Bonferroni correction was applied.

Results

Between 10:00 and 20:00 the HR slopes were higher in H, HS and HO groups than in the corresponding normotensive groups while in the two groups with dyslipidemia the values were quite similar. From 20:00 to 04:00, the HR slopes in NHS groups were higher than in NH while in NHD and in NHO the slopes presented lower values than in NH. In this period, the HS presented higher values of the slope than the other three hypertensive groups. In the early morning, the hypertensive groups presented higher slope values than normotensive groups showing that HR increased quicker in the hypertensive patients than in normotensive subjects (Figure 3.11 to 3.13, Table 3.6).

Figure 3.11: Circadian rhythms of HR in normotensive and hypertensive subjects with and without smoking (left panel) and their linear regression (right panel) in the three periods considered (10:00-20:00, 20:00-4:00, 5:00-10:00). NH=Normotensive (green points), NHS=Normotensive smoking (red points), H=Hypertensive (black points), HS=Hypertensive smoking (blue points).

Figure 3.12: Circadian rhythms of HR in normotensive and hypertensive subjects with and without dyslipidemia (left panel) and their linear regression (right panel) in the three periods considered (10:00-20:00, 20:00-4:00, 5:00-10:00). NH=Normotensive(green points), NHD=Normotensive with dyslipidemia (red points), H=Hypertensive (black points), HD=Hypertensive with dyslipidemia (blue points).
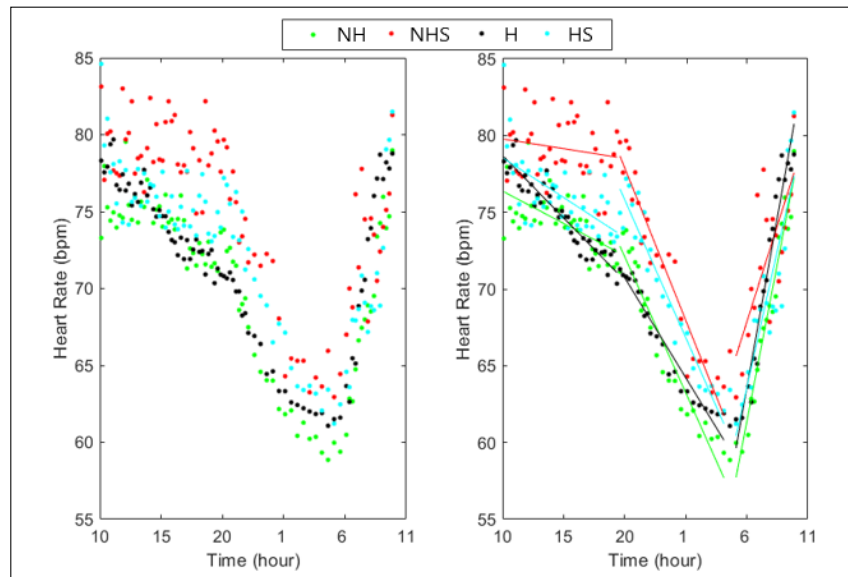

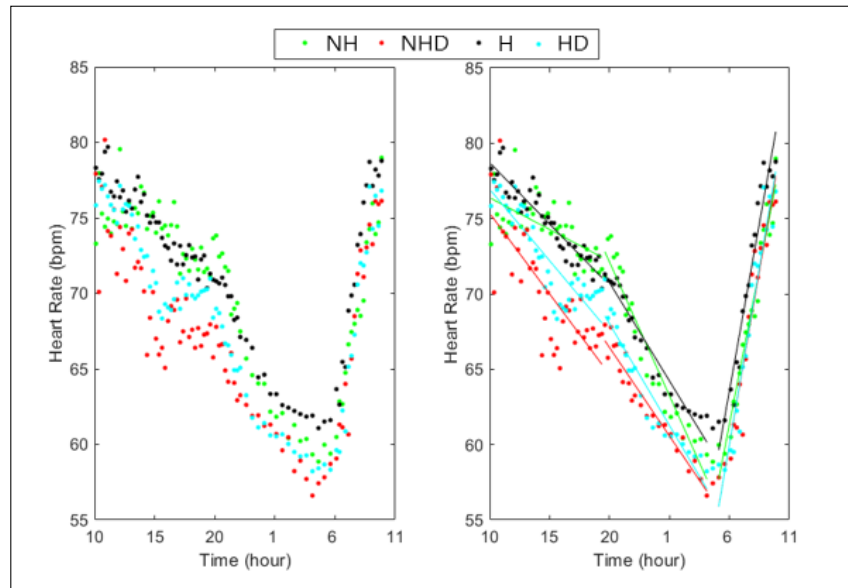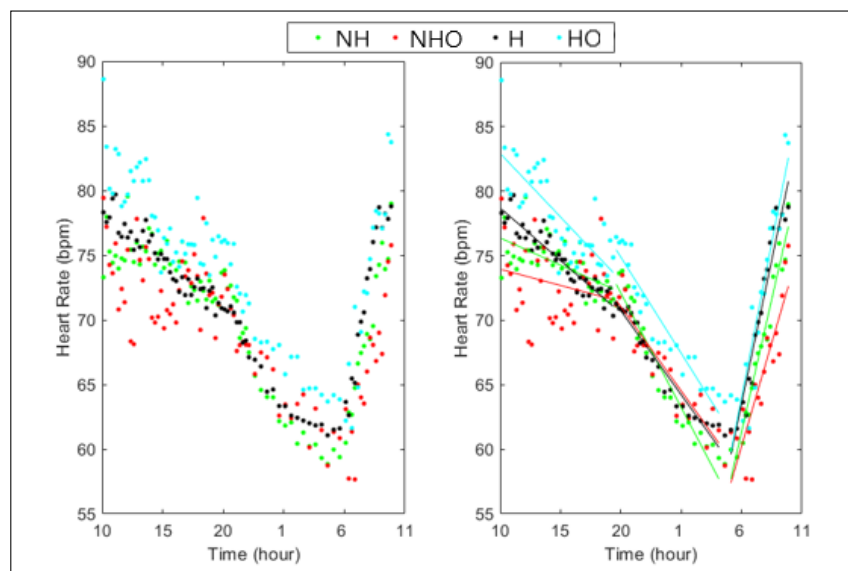**Figure 3.13:** Circadian rhythms of HR in normotensive and hypertensive subjects with and without obesity (left panel) and their linear regression (right panel) in the three periods considered (10:00-20:00, 20:00-4:00, 5:00-10:00). NH=Normotensive (green points), NHO=Normotensive obese (red points), H =Hypertensive (black points), HO=Hypertensive obese (blue points).

		NH	NHS	NHD	NHO	H	HS	HD	HO
Slope [bpm/hour]	10:00-20:00	-0.42	-0.13	-1.07	-0.25	-0.81	-0.52	-0.95	-0.99
	20:00-04:00	-1.77	-1.99	-1.17	-1.30	-1.30	-1.80	-1.35	-1.49
	05:00-10:00	4.12	2.52	4.61	3.21	4.45	3.57	4.68	4.80
Intercept [bpm]	10:00-20:00	80.6	81.06	86.06	76.48	86.86	83.76	86.43	92.80
	20:00-04:00	107.6	117.65	89.92	97.09	96.83	111.80	95.01	104.70
	05:00-10:00	-62.04	-7.59	-78.41	-35.92	-69.85	-43.50	-80.50	-79.92
R square	10:00-20:00	0.63	0.17	0.80	0.25	0.94	0.62	0.88	0.79
	20:00-04:00	0.96	0.96	0.97	0.93	0.98	0.97	0.96	0.95
	05:00-10:00	0.98	0.75	0.96	0.88	0.98	0.89	0.97	0.93

Table 3.6: Slope, intercept and R^2 values of the linear regression calculated in the three time periods in the eight subject groups.

In the first period, there were significant differences between each pair of groups except between H vs HS (Table 3.7-3.8). Between 20:00 and 04:00, there were significant differences only between NH vs NHS, H vs HD and H vs HO. The means along 24hrs shown significant differences as during the day time, except between NH and NHO (Table 3.8). Considering the comparison between NH and H groups presenting the same risk factor (Table 3.9), in the first period, as well as in day time, the differences were statistically significant only when a risk factor was present. In the interval 20:00-04:00 no significant difference was found between each pair while during awakening and night time only the difference between NHO and HO was significant. Along the 24hrs the differences between NH and H groups were significant only considering smoking and obesity (Table 3.9).

Discussion

During day time and early morning, H without other risk factors as well as smokers or presenting obesity showed greater velocity change than corresponding normotensive subjects. In the night time, the differences between the slopes in HO and in NHO as well as between HS and NHS patients became similar, while H patients showed lower slopes than NH subjects. In subjects affected by dyslipidemia, the differences between NH and H were negligible in all periods. The slopes calculated in the third period could be used to quantify the morning HR surge associated with acute cardiovascular effects while the values calculated from 20:00 to 04:00 could be used to measure the decline during night. The different slopes could indicate different control ways of HR changes along the 24hrs depending on the risk factor. Considering the mean HR values, the

results (Table 3.7-3.8) highlighted that all risk factors affected the mean HR values with a significant variation during day. Smokers presented significant higher HR values than non-smokers with NHS subjects showing the highest values and NHNS subjects the lowest. Higher values in NH subjects supporting the hypothesis that smoking has a larger impact on this category of subjects than in H confirming the outcomes reported by [36, 35] due to a worse oxygen exchange in the lungs of smokers. Considering the four hypertensive groups, the smoking does not affect significantly the HR mean values in any of the three periods. Hypertensive patients with obesity showed statistically higher HR values during all the three periods than hypertensive without risk factors, confirming the results of [39, 43] obtained on patients affected by metabolic syndrome during both day time and night time and in office condition. Moreover, the results suggested that the presence of dyslipidemia reduce the HR values during all day, significantly from 10:00 to 04:00, extending the results of [44, 45] in which only office measurements were considered. This reduction was probably due to the reduced activation of sympathetic nervous system. For hypertensive patients, the mean HR values calculated on day time were comparable to those evaluated during the first period (10:00-20:00) while the values estimated on 24hrs as well as on night time presented intermediated values between those of the second and third periods. In this study, it was highlighted that from 10:00 to 04:00 smokers presented higher HR values and subjects with dyslipidemia lower values than subjects without risk factors both in H and in NH subjects.

	NH	NHS	NHD	NHO	N	HS	HD	HO
10:00-20:00	74±2	79±2	70±4	73±3	75±2	76±2	72±3	78±3
20:00-04:00	66±5	71±5	62±3	66±4	66±3	70±5	63±3	70±4
05:00-10:00	68±6	72±5	68±7	66±5	71±6	70±6	68±7	72±7
Day time (08:00-21:00)	74±2	78±3	70±4	72±3	75±3	76±3	72±3	78±3
Night time (23:00-06:00)	61±2	66±3	59±2	63±2	63±1	65±3	60±1	66±2
24hrs	71±6	75±5	67±6	69±5	71±5	73±5	69±6	74±6

Table 3.7: Mean values ± SD of HR values (bpm) in the three periods, during day time, night time and along the 24hrs, for each subject group.

	NH vs NHS	NH vs NHD	NH vs NHO	H vs HS	H vs HD	H vs HO
10:00-20:00	<0.0006	<0.0006	0.02	n.s.	0.002	<0.0006
20:00-04:00	0.02	n.s.	n.s.	n.s.	0.04	0.04
05:00-10:00	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Day time (08:00-21:00)	<0.00006	<0.00006	0.002	n.s.	0.0006	<0.0006
Night time (23:00-06:00)	<0.00006	n.s.	n.s.	n.s.	<0.0006	<0.002
24hrs	<0.00006	0.001	n.s.	n.s.	0.009	0.01

Table 3.8: P-value of the comparison between normotensive and hypertensive groups with and without risk factors.

	NH vs H	NHS vs HS	NHD vs HD	NHO vs HO
10:00-20:00	n.s.	<0.0004	0.006	<0.0004
20:00-04:00	n.s.	n.s.	n.s.	n.s.
05:00-10:00	n.s.	n.s.	n.s.	0.03
Day time (08:00-21:00)	n.s.	<0.00004	0.002	<0.00004
Night time (23:00-06:00)	n.s.	n.s.	n.s.	0.01
24hrs	n.s.	0.001	n.s.	<0.00004

Table 3.9: P-value of the comparison between normotensive and hypertensive groups presenting the same risk factor.

Main finding

Until now, to evaluate the HR changes due to presence of these risk factors, a single HR office measure or a mean evaluated on day time or night time or 24hrs was used. However, since HR shows a circadian behavior with three characteristic behaviors, a single value represents only a rough approximation of this behavior. Applying linear regression analysis in each period, the slopes could be used to estimate the morning HR surge associated with acute cardiovascular effects in the awakening and to evaluate the decline during the night. Furthermore, considering the influence of the risk factors, the smoking increased and dyslipidemia decreased mean HR values from 10:00 to 04:00, both in normotensive and hypertensive subjects in comparison with subjects without risk factors. During the awakening (05:00-10:00) the slopes were similar among all groups with no significant difference among the mean HR values.

3.2.3 Influence of some cardiovascular risk factors on the relationship between Blood Pressure and age [7SG]

Materials and methods

Subjects

The study population included a total of 880 subjects (365 males and 524 females, 65±16 mean age). The information about risk factors were collected in accordance with international guidelines [143]. The subjects were then grouped in eight classes considering them either all together or excluding those presenting at least one risk factor, as reported in Table 3.10.

	All together		**Without risk factors**	
	Office	ABPM	Office	ABPM
H	253	241	54	60
NH	112	124	105	99

Table 3.10: Subjects group.

Statistic analysis

The SBP/Age and the DBP/Age relations were separately evaluated for H and NH subjects, by calculating the slope, the intercept and the R^2 of a regression line the parameters of the regression lines in each subject group.

Results

Considering subjects with and without risk factors, the intercepts were higher in H than in NH subjects and they were 2-8 mmHg greater in office than in ABPM measurements (Figure 3.14). The SBP/Age slopes were positive in both subject groups with higher values in office condition than in ABPM, while the DBP/Age slopes were negative in all subjects and slightly greater in ABPM than in office (Table 3.11). In all the subjects without risk factors (Figure 3.15), the SBP/Age intercepts were comparable in office and ABPM conditions while the DBP/Age intercepts were greater in hypertensive subjects in both conditions.

Figure 3.15: SBP/Age and DBP/Age relationships in the subject without risk factors (black: hypertensive patients; red: normotensive subjects) calculated in office condition (top panels) and as average on 24hrs (bottom panels)

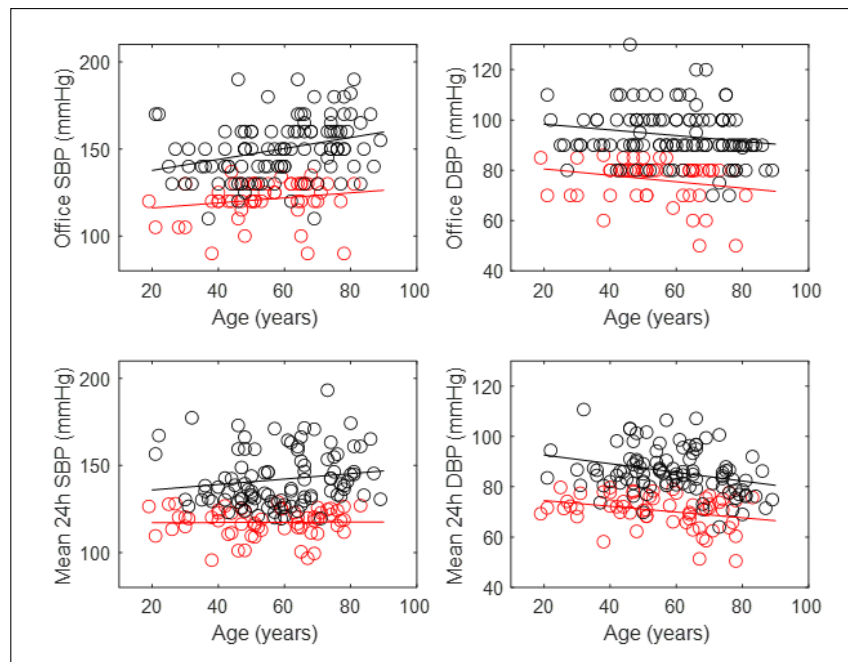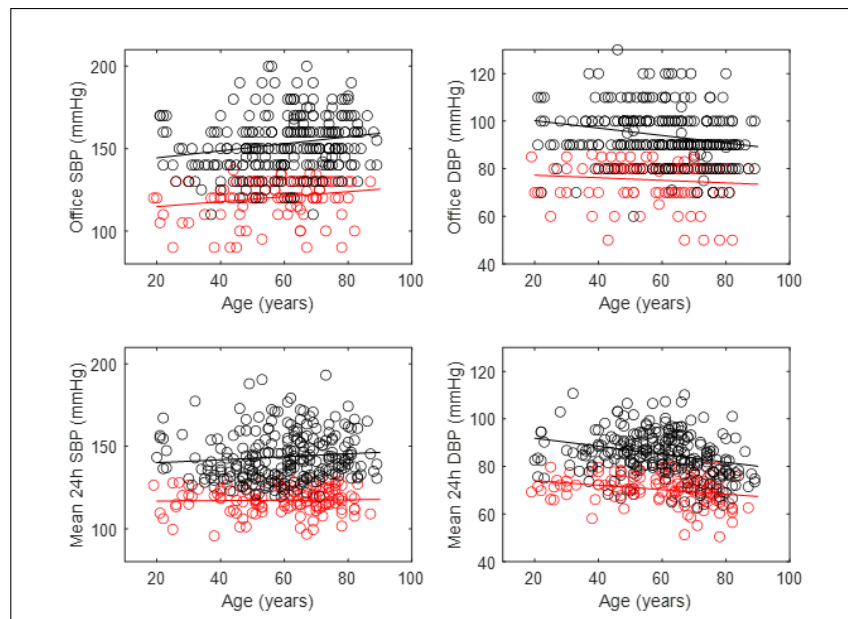**Figure 3.14:** SBP/Age and DBP/Age relationships in all the subjects (black: hypertensive patients; red: normotensive subjects) calculated in office condition (top panels) and as average on 24hrs (bottom panels)

All the subjects					
		Office		ABPM	
		NH	H	NH	H
SBP/Age	Intercept	112	140	116	138.
	Slope	0.15*	0.22*	0.01	0.09
DBP/Age	Intercept	78	103	76	95
	Slope	-0.05	-0.16*	-0.10*	-0.17*
Subjects without risk factors					
		Office		ABPM	
		NH	H	NH	H
SBP/Age	Intercept	113	131	117	133
	Slope	0.15	0.32*	0.005	0.157
DBP/Age	Intercept	83	101	77	96
	Slope	-0.13	-0.11	-0.12*	-0.17*

Table 3.11: Subjects group.

The SBP/Age slopes were positive in all measurements with the lowest values in normotensive subjects evaluated in ABPM, while the DBP/Age slopes presented lower values in office than in ABPM and in NH than in H subjects (Table 3.11). The differences between the regression lines of SBP/Age relation calculated on all subjects and on subjects without risk showed, in NH subjects (red lines in Figure 3.16, left panels), negligible and independent behavior of age while, in H subjects (black lines) the differences were related to age, decreasing toward zero in elderly. These trends were similar for BP calculated both in office and in ABPM ways. On the contrary, the differences between the regression lines of the DBP/Age relation decreased with ageing more remarkable if office was used, especially in NH subjects (red lines in Figure 3.16, right panels). In H the differences had trends that were opposite in office and ABPM, presenting a negligible relation with ageing. The differences between the regression lines of the SBP/Age relation calculated in office and ABPM ways in H (Figure 3.17, left panels) as well as in N subjects increased with age either in all the subjects or in those without risk factors. The differences between the regression lines of the DBP/Age relation increased with ageing when the subjects were considered all together while, if we considered only those without risk factors, the increase with age was presented only in H (Figure 3.17, right panels).
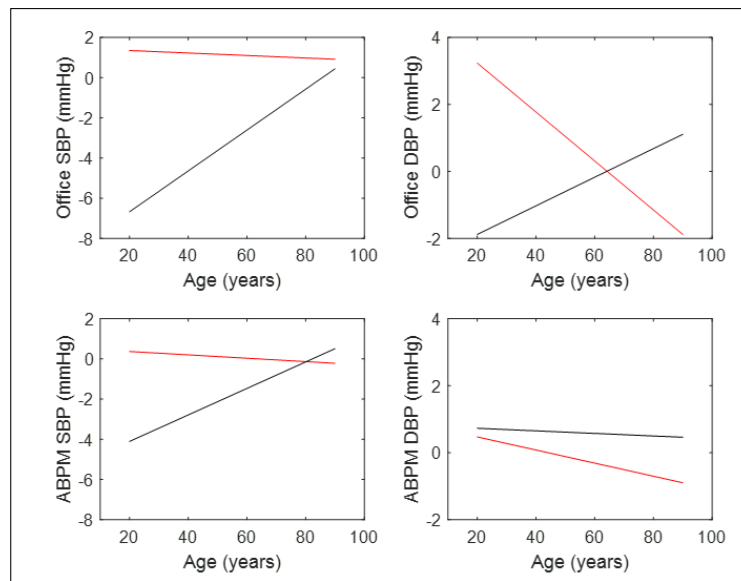
Figure 3.16: Regression lines of SBP/Age and DBP/Age relationships calculated on all the subjects and on subjects without risk factors, in office (top panels) and ABPM (bottom panels) modalities. Black lines: hypertensive patients; red lines: normotensive subjects.
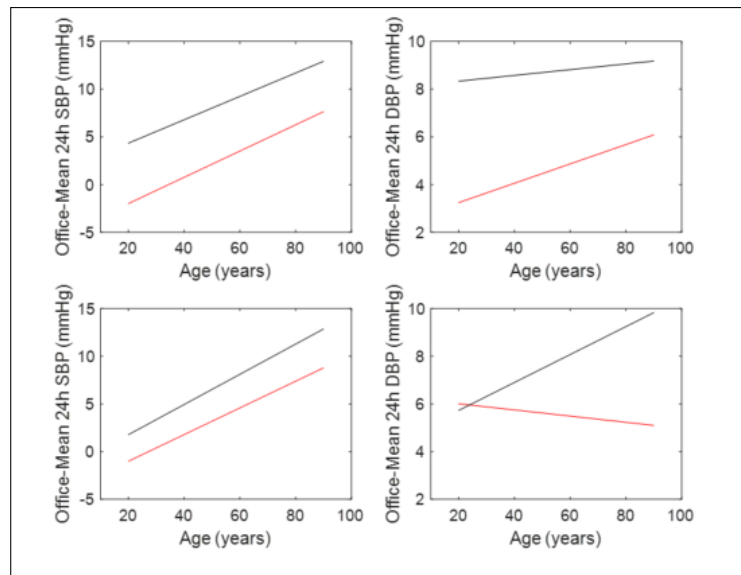

Figure 3.17: Differences between the regression lines of SBP/Age and DBP/Age relationships calculated in office and ABPM modalities on all the subjects (top panels) and on subjects without risk factors (bottom panels). Black lines: hypertensive patients; red lines: normotensive subjects.

Discussion

Several studies evaluated the relation between BP and age measurements but none of them distinguished between subjects with and without the presence of at least one risk factor although it is known that many factors could increase BP values [47, 48, 50]. Our results suggest that SBP increased significantly with age (Table 3.11), presenting different values (between 1.5 and 2.2mmHg per decade) in all subjects only for office measurements, while for ABPM the increment was negligible. On the other hand, DBP, after a quite constant behavior in the first decades of age, significantly decreases over 50 years, especially for BP measured in ABPM way. According to [54, 55, 56, 57] the increase in DBP with ageing maybe due to a large artery stiffness and to the peripheral vascular resistance in small vessel, instead the increase of DBP up to the age of 50 was due to peripheral vascular resistance and the successive decrease was due to the increase of large artery stiffness. About subjects without risk factors, our results highlighted that the dependence on age of DBP in H, both considering office and ABPM measures, was greater than that assessed considering all subjects together, with decreasing differences with age (black lines in Figure 3.16, left panels). The DBP/Age relationships were substantially the same considering all the subjects or only those without risk factors when ABPM was used, while differences of opposite sign in H and NH were present for the office measures (Figure 3.16 , right panels). Comparing the differences between the SBP/Age relations obtained measuring the BP in office and in ABPM (Figure 3.17), a large increase of the differences with age is present in all the situations. The differences in the DBP/Age relations increase with age, although less than in the SBP/Age except in normotensive subjects without risk factors for which the difference is approximately constant and equal to 5-6mmHg in all ages considered (Figure 3.17). Those findings highlighted the risks involved in estimating a correct BP when using the office measurement and underlines how ageing increases the error introduced by the measurement method in subject with or without risk factors.

Main findings

In subjects with and without risk factors (smoking, obesity and dyslipidemia), the slopes of the Systolic BP/Age relation were higher in hypertensive than in normotensive subjects in both office and ABPM conditions. This fact highlighted the risk involved in estimating a correct BP when using the office measurement and underlines how ageing increases the error introduced by the measurement method in subject with or without risk factors.

3.2.4 Influence of ageing on circadian rhythm of HRV in normal subjects [8SG]

Materials and methods

Subjects

A total of 140 (70 men and 70 female) normal subjects aged from 15 to 90 years separated into three age groups: 15-39 years old (47 subjects, Young Group (YG)); 40-64 years old (47 Adult Group (AG)); 65-90 years old (46 Senior Group (SG)).

Statistical analysis

Several linear and non-linear HRV parameters were calculated on each segment and they were averaged among the subjects of each age group. The HRV parameters were examined averaging (for each subject and age group) the values in three different time periods: along the whole 24hrs, during the day (from 14:00 to 19:00) and during the night (from 2:00 to 7:00). Finally, the relationship between age and each parameter, calculated on each subject as the average along the 24hrs, was considered and modelled by using linear and quadratic approximations. Adjusted determination coefficient, R^2, was used to compare the two models. The comparisons between groups were carried out by the Wilcoxon rank sum test, assuming that the subject groups were independent and showed a non-Gaussian distribution, and Bonferroni correction was conducted. The Brown–Forsythe statistic test or the non-parametric Kruskal-Wallis test were used when the assumption of equal variances did not hold. The difference between each pair of groups was considered significant if p-value was lower than 0.05.

Results

Circadian rhythm of the MeanRR was comparable between groups (Figure 3.18) highlighting that the YG showed the greatest difference between day and night mean values while SG presented the lowest one. Moreover, the AG presented intermediate mean values with a maximum value during the night, similar to the SG, and a minimum value during the day, comparable to the YG (Table 3.12). Significant differences of these values between each pair of groups (Table 3.13) were found only during day time for AG vs SG (p<0.02).

	YG				AG				SG			
	24h	Day	Night	Day vs Night	24h	Day	Night	Day vs Night	24h	Day	Night	Day vs Night
MeanRR (ms)	868±138	795±83	1015±68	<0.00001	846±120	771±70	958±71	<0.00001	884±92	840±66	952±59	<0.00001
SDNN (ms)	85±33	79±24	100±35	<0.0001	60±26	56±20	65±28	0.006	57±25	58±21	61±25	ns
NN50	77±39	66±29	101±29	<0.0001	32±24	28±17	38±20	0.007	35±25	40±19	33±20	ns
pNN50	0.24±0.14	0.19±0.09	0.35±0.11	<0.0001	0.10±0.08	0.08±0.05	0.13±0.07	0.0007	0.11±0.07	0.12±0.06	0.11±0.07	ns
RMSSD (ms)	67±28	58±20	86±23	<0.0001	46±24	44±19	48±20	ns	56±28	57±23	58±26	ns
LF (ms2)	1984±1514	1788±1041	2498±1810	0.0001	998±945	950±701	1099±891	0.02	887±902	837±702	943±867	ns
HF (ms2)	1709±1347	1227±894	2554±1217	<0.0001	716±810	672±628	761±614	ns	1116±1148	1124±870	1200±1046	ns
LFn	0.62±0.15	0.66±0.13	0.55±0.14	<0.0001	0.63±0.17	0.65±0.14	0.59±0.16	0.001	0.52±0.16	0.52±0.14	0.53±0.16	ns
HFn	0.38±0.15	0.34±0.13	0.45±0.14	<0.0001	0.37±0.17	0.35±0.14	0.41±0.16	0.001	0.48±0.16	0.48±0.14	0.47±0.16	ns
LF/HF	2.6±2.0	3.0±1.8	2.0±1.4	<0.0001	2.9±2.4	3.3±2.3	2.4±2.0	0.002	1.9±1.7	1.7±1.2	2.1±1.7	ns
SD1/SD2	0.23±0.08	0.22±0.06	0.24±0.07	ns	0.24±0.09	0.24±0.08	0.22±0.08	0.02	0.28±0.11	0.30±0.10	0.26±0.09	0.002
Beta exponent	1.27±0.35	1.30±0.30	1.25±0.34	ns	1.28±0.40	1.26±0.36	1.31±0.40	ns	1.00±0.42	0.96±0.36	1.06±0.41	0.02
FD	1.45±0.11	1.44±0.09	1.47±0.09	ns	1.47±011	1.48±0.10	1.46±0.09	ns	1.56±0.11	1.59±0.09	1.52±0.09	0.0001

Table 3.12: Mean values (±SD) of all the parameters in the three groups during 24hrs, day and night times together with p-values of the differences between day and night in each group.
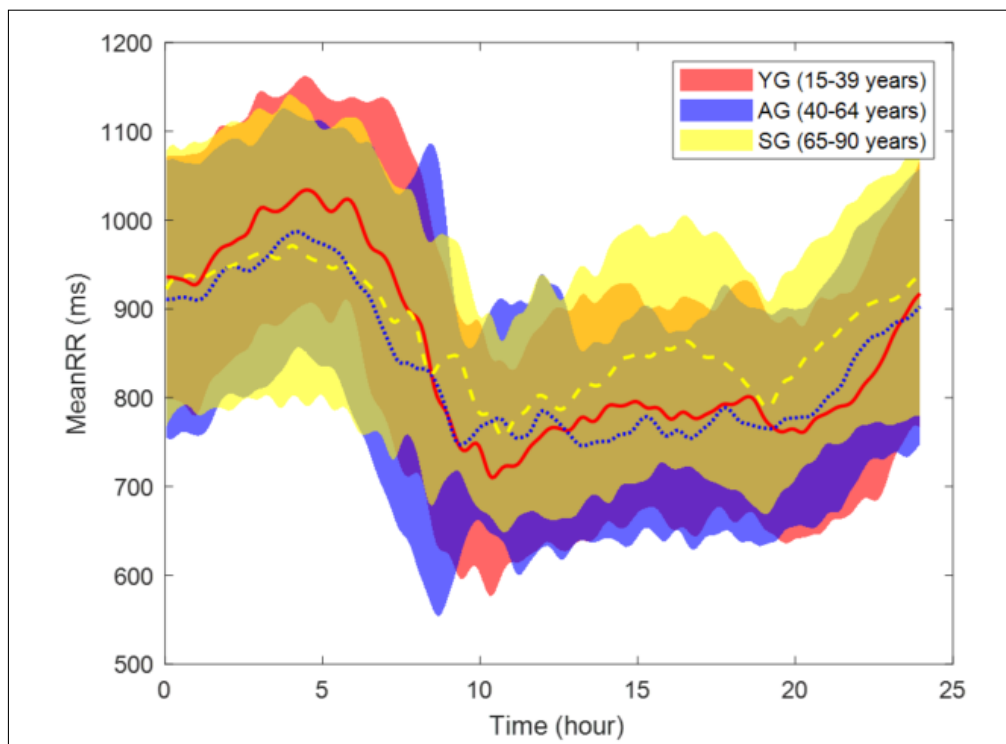

Figure 3.18: Circadian rhythm of MeanRR values in the three age groups together with ±1SD.

	24h			Day			Night		
	YGvsAG	YGvsSG	AGvsSG	YGvsAG	YGvsSG	AGvsSG	YGvsAG	YGvsSG	AGvsSG
MeanRR (ms)	ns	ns	ns	ns	ns	0.02	ns	ns	ns
SDNN (ms)	<0.0001	<0.0001	ns	<0.0001	0.0001	ns	<0.0001	<0.0001	ns
NN50	<0.0001	<0.0001	ns	<0.0001	0.002	ns	<0.0001	<0.0001	ns
pNN50	<0.0001	<0.0001	ns	<0.0001	0.004	ns	<0.0001	<0.0001	ns
RMSSD (ms)	0.03	ns	ns	0.03	ns	ns	<0.0001	0.002	ns
LF (ms2)	<0.0001	<0.0001	ns	0.0001	<0.0001	ns	<0.0001	<0.0001	ns
HF (ms2)	0.0007	0.02	ns	0.003	ns	ns	<0.0001	0.001	ns
LFn	ns	0.002	0.0003	ns	<0.0001	0.0001	ns	ns	0.05
HFn	ns	0.002	0.0003	ns	<0.0001	0.0001	ns	ns	0.05
LF/HF	ns	0.007	0.0002	ns	0.0002	0.0001	ns	ns	ns
SD1/SD2	ns	0.01	0.03	ns	0.002	ns	ns	ns	ns
Beta exponent	ns	0.0008	0.001	ns	0.0001	0.002	ns	ns	0.008
FD	ns	<0.0001	0.0001	ns	<0.0001	0.0001	ns	0.03	0.01

Table 3.13: P-values of the differences between each pair of groups for all the parameters averaged on 24hrs, day and night times.

For all parameters calculated in the time domain, the YG showed a sudden and significant increase between day and night time, after 20:00, and a progressive decrease after 7:00 (Figure 3.19, Table 3.12 - 3.13). Considering the frequency domain, both YG and AG had significantly greater values of LF/HF and LFn and significantly lower values ($p<0.002$) for HFn compared to SG ($p<0.01$), while during night time these differences were negligible (Figure 3.20, Tables 3.12 - 3.13). LF and HF values showed a similar trend in all three groups and significant differences between day and night time were found only in YG (Figure 3.20, Table 3.11). Furthermore, for parameters averaged along the 24hrs, the YG showed significant differences compared to both AG and SG groups (Table 3.13). Regarding the three non-linear parameters, significant differences ($p<0.02$) between night and day were found for all parameters in SG and for SD1/SD2 in AG (Table 3.12). Furthermore, SG showed significant differences ($p<0.03$) with respect to the YG and AG groups during 24hrs and day time periods for all the three parameters (Table 3.13, Figure 3.20). During night time, these differences remained significant ($p<0.03$) only for FD and Beta exponent parameters (Table 3.13, Figure 3.20). Figure 3.21 shows the relationships between age and those parameters that highlighted a parabolic-like behaviour, suitably modelled by a quadratic relation. In particular, LF/HF, LFn and Beta exponent presented an increasing trend until 55-60 years and a decreasing trend over that age. Instead, SDNN, NN50, pNN50, RMSSD, SD1/SD2 and HFn showed an opposite behaviour. For these parameters, the adjusted determination coefficient, R^2, showed greater values (at least of 25%) for quadratic regression than for a linear one.
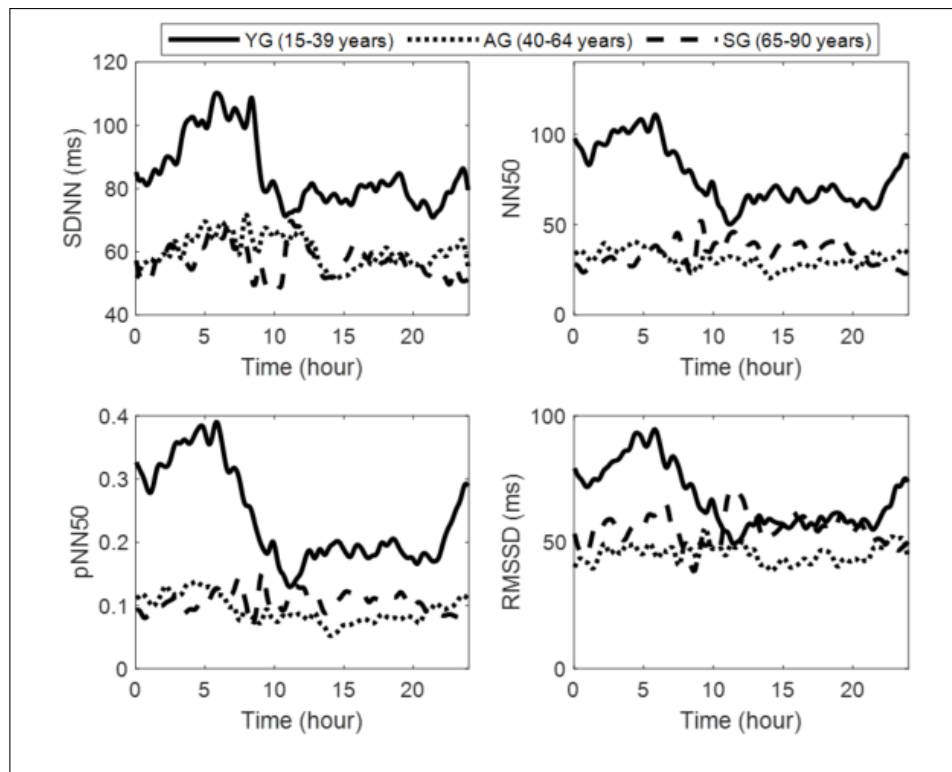
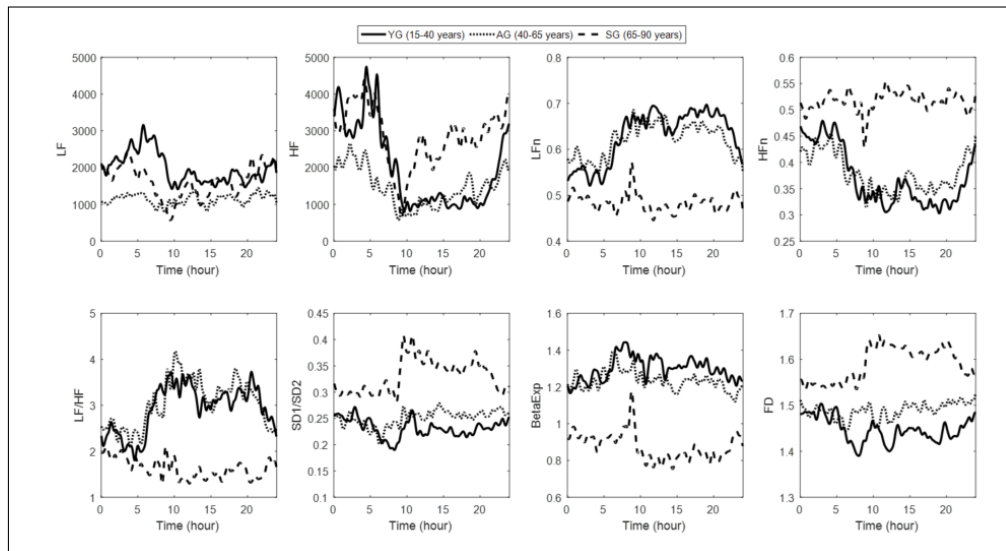**Figure 3.19:** Circadian rhythm of SDNN, NN50, pNN50 and RMSSD linear parameters in the three age groups.

Figure 3.20: Circadian rhythms of LF, HF, LFn, HFn, LF/HF, Beta exponent, SD1/SD2 and FD parameters in the three age groups.
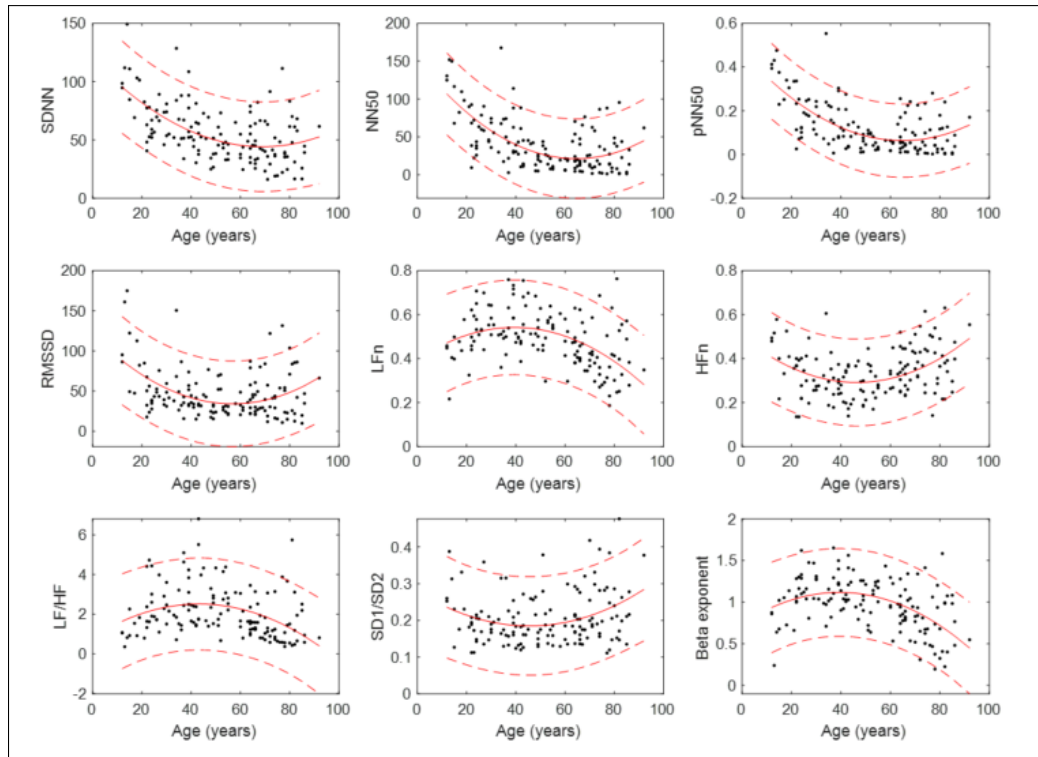
Figure 3.21: Quadratic relationships between linear and non-linear parameters and age calculated considering the average among the 24hrs of each subjects. Solid lines represent the parabolic relation and dashed lines represent 95% confidence intervals.

Discussion

The differences of MeanRR (Table 3.12) reflected a higher vagal tone in the young subjects, probably blunted in the oldest group (Figure 3.18), probably due to a better physiological adaptation in younger subjects. The YG presented significant higher values of SDNN, NN50, pNN50 and RMSSD during the night than during the day (Table 3.12), which decreased with aging with significant difference between YG and the other two groups (Table 3.13), suggesting a reduced autonomic regulation of heart with aging [60, 72]. The oldest group were characterized by lower fluctuations between day and night (Table 3.12), highlighting a possible deterioration of cardiac autonomic nervous control over time. The spectral parameters were very similar for AG and YG (Figure 3.20 and Table 3.13), presenting significant differences between night and day, as previous reported [43]. On the contrary, the oldest group had flattened variation between day and night (Table 3.12). These results confirm a loss of variability and complexity already found in subjects in whom age increased up to 64 years [73]. The FD values did not present significant differences between day and night for subjects younger than 64 years (Table 3.12). In older subjects, these parameters showed significant differences between day and night and in the mean values over 24hrs, compared to the others two

groups (Table 3.13). These findings highlighted a more complex generating system in SG compared to the YG and AG groups, probably due to a higher variability present in older subjects. A specific parabolic age-related shaped analysis was found between almost all the considered parameters and age. In particular, SDNN, NN50, pNN50, RMSSD, HFn, SD1/SD2 parameters showed a negative correlation with ageing from 15 to 60 years and an inverse trend after the age of 60, while LFn, LF/HF and Beta exponent presented an opposite behaviour (Figure 3.21). MeanRR, LF and HF parameters showed a simple linear decreasing trend with aging without any evident behaviour change in elderly. The inverse behaviour presented in elderly over 75 years compared to young and adult subjects could be explained by a decreased compensatory response of the RR variability during the 24hrs in the elderly. This important findings underline the role of autonomic imbalance and decreased compensatory response with time. The value of HRV in physiological aging highlights its clinical applicability to predict autonomic response in older subjects.

Main findings

This is the first study in literature that examined changes in linear ad non-linear HRV parameters in subjects over 65 years old using an accurate short-term evaluation on 288 consecutive intervals of 5 minutes among the 24hrs. The subjects were divided in tree age groups: young group 15-39 years old; adult group 40-64 years old; senior group, 65-90 years old. The time domain parameters (MeanRR, SDNN, NN50, pNN50, RMSSD) presented significant differences between young group and the other two groups along the 24hrs while normalized spectral parameters (LFn, HFn) and non-linear parameters (SD1/SD2, Beta exponent and FD) showed significant differences between senior group and the other two groups. Evaluating the circadian rhythm of all parameters, a significant difference between mean day and night values was found. Another remarkable result was the parabolic-shaped relationship between each parameter and age highlighting an opposite trend over about 60 years old compared to younger sign of a progressive physiological autonomic imbalance with ageing.

3.2.5 Influence of the gender on the relationship between Heart Rate and Blood Pressure [9SG]

Materials and methods

Subjects

The study population consisted of 172 normotensive subjects was com-posed of 50 males (aged 61±17) and 122 female (aged 57± 19).

Statistical analysis

To examine the circadian trend of the relationship between BP and HR, the values of these parameters were averaged hour-to-hour in each subject. The slopes

and the intercepts of the regression lines fitting the relationships SBP/HR and DBP/HR for each gender were calculated. The significance of each relationship was evaluated by examining the p-value of the difference between the slope and zero. In order to compare the results with the literature, the relationship BP/HR was also evaluated considering, for each subject, either the office measurements or Mean average over 24hrs (ABPMm).

Results

SBP/HR presented slope values significantly different from zero (p-value < 0.05) from 14:00 to 18:00, from 21:00 to 1:00 and from 6:00 to 12:00 in females and only for few hours in males (Figure 3.22). The DBP/HR relation showed slopes significantly different from zero (p-value<0.05) in all hours of the day for females and from 6:00 to 17:00, except at 7:00 and 10:00, for males. The slope values were comparable in male and female groups in both relationships from 10:00 to 21:00 while after that the males' slopes decreased. The intercepts showed similar values between the two subject groups during day time; while during the night females had lower values (Figure 3.23). Figure 3.24 shows the mean values of SBP, DBP and HR caluclated hour-to-hour on females and males separately. The circadian trend in the two subject groups was similar with lower values in the night (Figure 3.24), females showed lower values for BPs than males along the 24hrs. On the contrary, HR values were higher in females than males during day time and similar during night time.
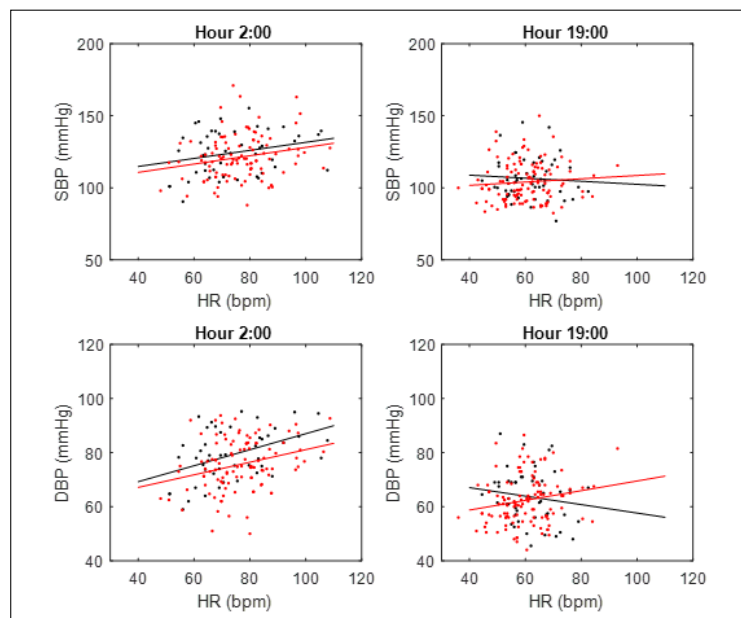
Figure 3.22: Examples of the relationship between either SBP (top panels) or DBP (bottom panels) and HR at two different hours during 24hrs, for both male (black) and female (red) groups together with corresponding regression lines. Left panels at 2:00; Right panels at 19:00.
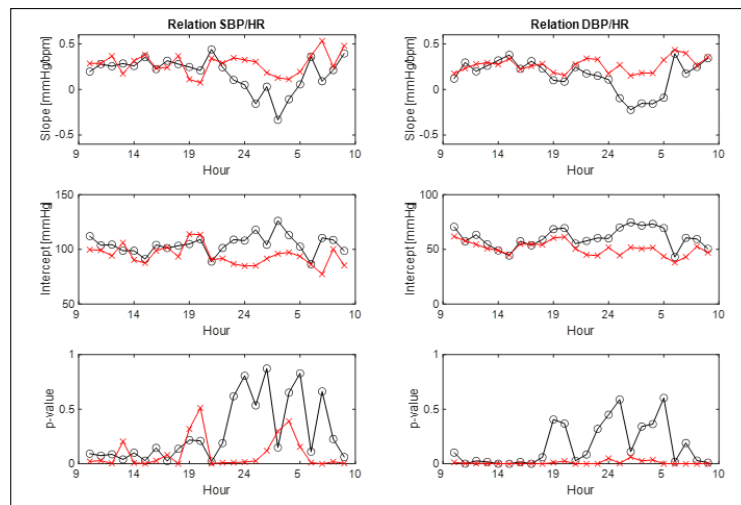
Figure 3.23: 24hrs behaviors of the slopes, intercepts and p-values of the hour-to-hour regression lines of the SBP/HR (left panel) and DBP/HR (right panel) relations in male (black circles) and female (red crosses) groups.
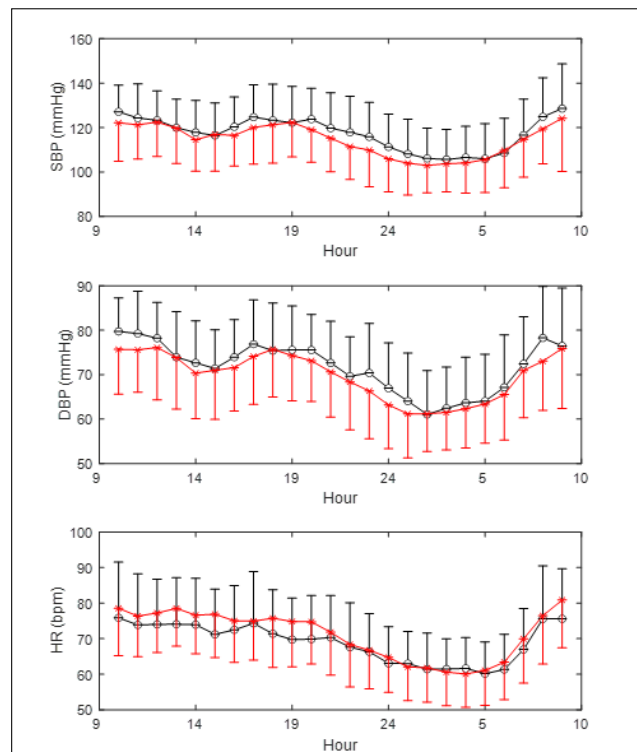
Figure 3.24: Mean values (±SD) of SBP, DBP and HR calculated hour-to-hour on females (red crosses) and males (black circles).

Figure 3.25: SBP/HR and DBP/HR relationships in females (red) and males (black) evaluated in office condition (left panels) and as ABPMm (right panels).

Figure 3.25 shows the relations SBP/HR and DBP/HR calculated in male and female subjects when BP and HR were evaluated either in office condition or as ABPMm. The correspondign slopes, intercepts and p-values are reported in Table 3.14. For SBP/HR relation (Figure 3.25) evaluated in office condition, the slopes were positive for males and negative for f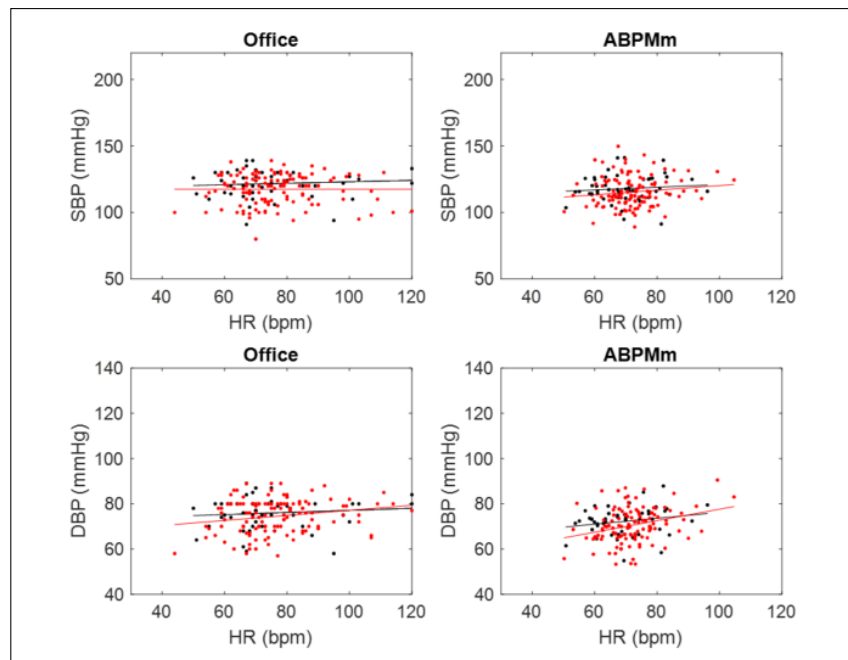emales and they were not significantly different from zero (Table 3.14). Considering the ABPMm, females presented not significant higher slope values than males. Moreover, the intercepts in office condition were similar between genders and in ABPMm were higher in male than in female subjects. The slopes of the DBP/HR relation presented positive values in both subject groups with significant higher slopes in females than in males both in office and ABPMm. Figure 3.26 shows the behavior of BP/HR relations hour-to-hour along 24hrs in both subject groups. After 22:00 a marked decrease of HR and BP until 2:00 was present; successively from 5:00 to 10:00 both signals quickly increased presenting an about linear trend with high SBP/HR slopes.

Discussion

The analysis of BP/HR relationship in office highlighted a significant linear regression only for DBP in females (Figure 3.25 and Table 3.14) partially confirming from the literature [24]. In males, neither SBP nor DBP showed a significant relation with HR probably due to a large variability among our subjects. The BP/HR relation estimated by using the ABPMm values confirmed the results obtained in office. In females, the SBP/HR relation was significant

only during day time, partially justifying the results found in the office and on ABPMm [23, 66, 67]. The DBP/HR relationship was not significant in males for half a day while in females the relation was significant at all hours of the day, confirming what was found in the office and on ABPMm [23, 66, 67]. During the night, females maintained the DBP/HR relation fairly unchanged while males between 24:00 and 5:00 reverse this reation. This behavior suggest the association between HR and cardiovascular mortality presented only in females [23] that could be due to an increase in neural activity and to estrogen related vascular regulation that might predispose to ventricular fibrillation and sudden death [68]. The pair of HR and BP circadian values averaged on all subjects (Figure 3.26) showed how this relationship presented a similar trend in both gender groups. Moreover, two different behaviors during day time and night time with limited change of HR (5 bpm) and BP (10 mmHg) during day and quick variation of both variables (about 15-20 mmHg and about 20 bpm) during the night were reported. In conclusion, the study showed in females a more uniform behavior throughout night and day and a significant relationship between DBP and HR while in males the relation was quite not significant and remarkably different during the night in respect of the day time.

SBP/HR	m	q	p-value
M Office	0.06	118	n.s.
M ABPMm	0.10	111	n.s.
F Office	-0.0006	118	n.s.
F ABPMm	0.17	103	n.s.
DBP/HR	m	q	p-value
M Office	0.05	72	n.s.
M ABPMm	0.13	63	n.s.
F Office	0.11	66	0.03
F ABPMm	0.25	55	0.001

Table 3.14: Slopes (m), intercepts (q) and p-values of the linear fitting of SBP/HR and DBP/HR in office and as ABPMm, for male M and female F subjects. n.s.: not significant.
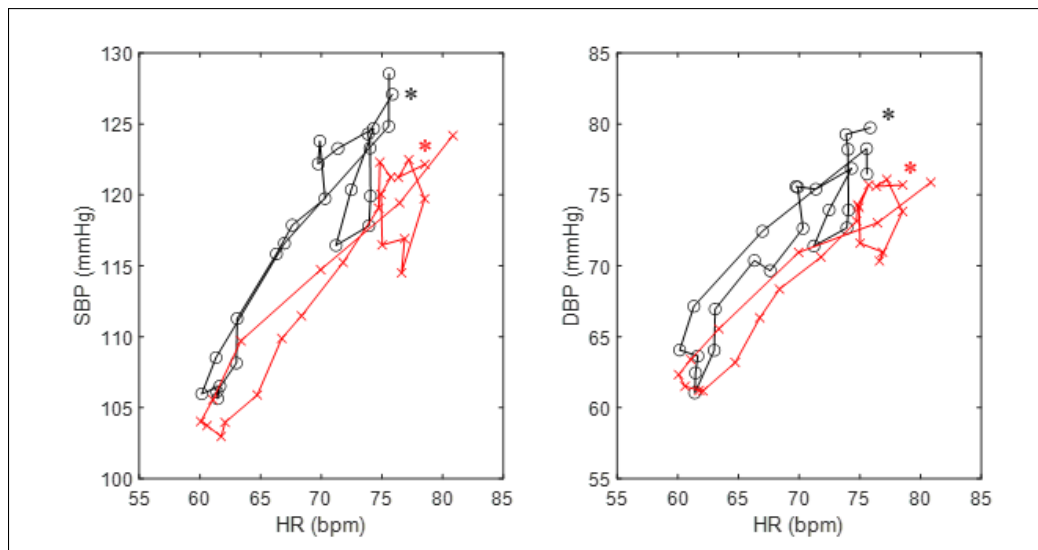
Figure 3.26: SBP/HR and DBP/HR relationships in females (red crosses) and males (black circles) evaluated hour-to-hour during 24hrs. Each cross/circle represents the values at a single hour. Asterisk indicates the starting time of each rhythm (10:00)

Main Findings

The circadian rhythm of BP/HR relationship was similar in both genders but with different values of HR and BP at different times of day and a different trend between genders only during night time.

3.3 Artificial Intelligence for identification of cardiovascular diseases

In this section, two different machine learning techniques were developed for identifying different pathologies such as ischemic heart (IHD) and dilated cardiomyopathy (DCM) diseases. Moreover, it was evaluated if adding a clinical parameter (LVEF) would present an increase of the classification performance. Initially, the artificial neural network (ANN) based on generalized back-propagation algorithm was developed to identify IHD patients. At first, several ANNs using as inputs HRV parameters, age, gender and LVEF were trained and validated in a large dataset. Successively, a feature selection was applied to reduce the inputs of ANNs but preserving the variance. Finally, several ANNs based on a combination of HRV parameters together with age and gender were evaluated and validated on a dataset matched by age and gender, in such way these factors will not affect the results. Successively, the Classification and Regression Tree (CART) was applied in order to identify DCM patients having as features the same input parameters. Finally, the CART method, considering the same input selection, was used to distinguish

the two different pathologies (IHD and DCM) from normal situation, on a large dataset.

3.3.1 ANN for Ischemic Heart Disease [10SG] [11SG] [12SG]

Materials and methods

Subjects

In the first study, the population consisted of 156 (80 male and 76 female, aged 53 ± 21) normal subjects and 87 (72 male and 15 female, aged 71 ± 10) suffered from ischemic heart disease. In the second paper, the cohort of subjects was composed of 681 (316 males, aged 62 ± 15 and 365 females, aged 64 ± 15) normal and 284 (222 males, aged 71 ± 10 and 62 females, aged 76 ± 10) subjects suffered from IHD. In the last paper, the study was composed of 496 subjects consisting of 251 normal subjects (188 males aged 73 ± 8 and 63 females aged 77 ± 7) and 245 suffered from IHD (186 males 74 ± 7 and 59 females aged 78 ± 7) matched by gender and age.

HRV parameters

Several linear (in time and space) and non-linear HRV parameters were evaluated on each segment and subject.

Neural Network Classifier

A multi-layer feed forward neural network with sigmoid activation function was developed and a series of networks with a different number of input nodes, hidden nodes and two output nodes were tested.

In the first study, the ANN presented as inputs the HRV parameters (MeanRR, SDNN, RMSSD, NN50, pNN50, LF, HF, LF/HF, LFn, HFn, Beta Exponent, SD1, SD2, SD1/SD2, FD), age, gender and LVEF.

In the second paper, three different combinations of parameters were considered as inputs of ANN. In the first situation (ANNa) the first principal components explaining at least the 90% of system variability was considered, in the second case (ANNb) the parameters extracted by using stepwise analysis presenting significant p-values ($p<0.05$) were used and finally, in the third situation, utilized for comparison, all the 15 linear and non-linear parameters were considered (ANNc) with age and gender. The ANNs characteristics were reported in Table 3.15.

	#input	#hidden neurons	Input parameters
ANNa	5	4	1°Comp, 2°Comp, 3°Comp, 4°Comp, 5°Comp
ANNb	5	2	MeanRR, LFn, SD1, age, gender
ANNc	17	3	MeanRR, SDNN, RMSSD, NN50, pNN50, LF, HF, LF/HF, LFn, HFn, Beta exponent, SD1, SD2, SD1/SD2, FD, age, gender

Table 3.15: ANNa, ANNb, ANNc characteristics.

Finally, in the last paper several combinations of parameters were considered. At first, only the linear parameters (ANN1) or only the non-linear parameters (ANN3) and all the fifteen parameters all together (ANN5) were considered as input. Successively, to reduce the number of input parameters, the Pearson correlation coefficient r among all parameters was evaluated and other two ANNs excluding either linear (ANN2) or non-linear (ANN4) parameters significantly ($p < 0.05$) correlated for more than 90% were examined. In addition, all the ANNs obtained selecting as input one or more linear parameters were considered for ANN2 combined together with one or more non-linear parameters considered for ANN4. We examined 31713 combinations of different network structure varying the number of inputs, equal to the number of the considered parameters, and of hidden neurons from 2 to 7. The input and hidden neurons combination producing the highest accuracy was then selected (ANN6). Finally, the LVEF was added to the inputs of each of the first five ANNs, considering ANN7 to ANN11 new networks. As the last case (ANN12), the procedure used for the ANN6 was repeated, by examining 31713 new ANNs varying, for each of them, the number of hidden neurons from 2 to 7, selecting the ANN producing the highest ACC.The features, the number of inputs and the number of hidden neurons selected for each ANN are reported in Table 3.16. In all papers, the training and test sizes were respectively 75% and 25% of the total number of data. The training of the neural network ended if the sum of the square errors for all data was less than 0.05 or the maximum number of training epochs (1000 iterations) was reached. Each of the twelve ANNs was simulated 100 times changing randomly the training and test dataset and the performance parameters were evaluated.

	#input	#hidden neurons	Input parameters
ANN1	12	3	MeanRR, SDNN, RMSSD, NN50, pNN50, LF, HF, LF/HF, LFn, HFn, age, gender
ANN2	10	2	MeanRR, RMSSD, NN50, pNN50, LF, HF, LF/HF, LFn, age, gender
ANN3	7	3	Beta exponent, SD1, SD2, SD1/SD2, FD, age, gender
ANN4	5	3	Beta exponent, SD2, FD, age, gender
ANN5	17	3	MeanRR, SDNN, RMSSD, NN50, pNN50, LF, HF, LF/HF, LFn, HFn, Beta exponent, SD1,SD2, SD1/SD2, FD, age, gender
ANN6	12	3	MeanRR, SDNN, NN50, pNN50, LF, HF, LF/HF, LFn, Beta exponent, SD2, age, gender
ANN12	8	4	MeanRR, LF, LF/HF, Beta exponent, SD2, age, gender, LVEF

Table 3.16: ANN characteristics.

Results

In the first study, to evaluate which network structure presented the highest accuracy, 135 combinations of different network structures varying the number of inputs (equal to the number of the considered parameters) from 4 to 18 and of the hidden nodes between 2 and 10 were examined. The structure with highest ACC, presented one hidden layer with 7 nodes and the combination of LVEF, gender, age, SDNN, pNN50, Beta exponent and SD2 as inputs. In the training phase, the accuracy was of 98.9% and the error rate was of 1.1% while using the validation dataset the accuracy was of 82% with an error rate of 18%. The area under the ROC curve was 0.99 for the training dataset and 0.83 for the validation dataset.

In the second paper, carrying out the principal component analysis, five most significant principal components, which account for 92% of the variance of the dataset, were selected as inputs for the first ANN (ANNa). Applying step-wise method, five features (MeanRR, LFn, SD1, gender and age) were used as inputs for the second ANN (ANNb). Finally, all the seventeen parameters were considered as input for the third artificial neural network (ANNc). The ANNs showing the highest accuracy in the three situations were those in which 4, 2 and 3 hidden neurons, respectively, were used. Among the three different ANNs, ANNc presented the lowest performance values while the other two classifiers had similar values with slightly higher values in ANNb (ACC max=82%) (Table 3.17).

		ANNa	**ANNb**	**ANNc**
SEN (%)	Max	52	58	49
	Mean±SD	46.6±5.9	47.5±9.5	44.1±5.9
SPE (%)	Max	92	90	87
	Mean±SD	46.0±3.1	86.3±3.6	84.4±3.4
PRE (%)	Max	74	69	61
	Mean±SD	58.6±6.1	59.6±5.3	54.1±6.3
ACC (%)	Max	80	82	77
	Mean±SD	74.4±2.5	74.7±2.6	72.5±2.5
AUC (%)	Max	83	85	80
	Mean±SD	78.1±2.8	77.78± 7.5	74.0± 3.4

Table 3.17: Mean (±SD) and maxima values of the classification performance indexes for the three ANNs in the test phase. SEN: sensitivity, SPE: specificity, PRE: precision, ACC: accuracy, AUC: area under the ROC curve.

In the third study, among 100 possible combinations of data used as input in the test for each ANN, it was selected that producing the highest ACC (Table 3.18). Almost all the best ANNs used 3 hidden neurons, only ANN8 and ANN12 used either 2 or 4 neurons, respectively. The performances obtained using LVEF were always better than those without it, achieving an accuracy of 79.8% with ANN12. The distribution values of the classification performance obtained using 100 different combinations of data input for the test set presenting similar quite symmetrical distributions. The mean (±SD) values of the sensitivity, specificity, precision and accuracy distributions for ANN2 and ANN12 were 67.8±8, 55.7±7, 59.2±5, 60.5±4 and 65.4±7, 75.1±7, 72.0±6, 70.2±4, respectively. Slightly better results were obtained in the training phase for all the ANNs. The mean (±SD) values for ANN2, without LVEF, were of 71.7±2 for training and 64.6±4 for test phase, respectively and for ANN12, with LVEF, of 86.1±2 and 77.5±4 in the two phases. Finally, Figure 3.27 reports the ROC curves for the ANN2 and ANN12, presenting the best accuracy, during training and test phases.

ANN Classifier	H	SEN%	SPE %	PRE %	ACC %	AUC %
ANN1	3	71.2	67.2	71.2	69.4	70.7
ANN2	2	80.3	62.1	70.7	71.8	71.5
ANN3	3	75.4	65.7	65.2	70.2	70.5
ANN4	3	74.2	58.6	67.1	66.9	73.7
ANN5	3	75.0	60.9	64.3	67.7	70.0
ANN6	3	83.1	56.9	63.6	69.4	74.5
ANN7	3	79.2	80.3	75.0	79.8	84.5
ANN8	2	80.0	78.1	77.4	79.0	84.7
ANN9	3	74.2	79.0	78.0	76.6	83.7
ANN10	3	69.6	83.8	78.0	77.4	84.7
ANN11	3	71.4	82.0	80.4	76.6	87.1
ANN12	4	75.4	84.1	82.1	79.8	86.0

Table 3.18: Best classification performance of ANNs; H=Hidden neurons, SEN: sensitivity, SPE: specificity, PRE: precision, ACC: accuracy, AUC: area under the ROC.
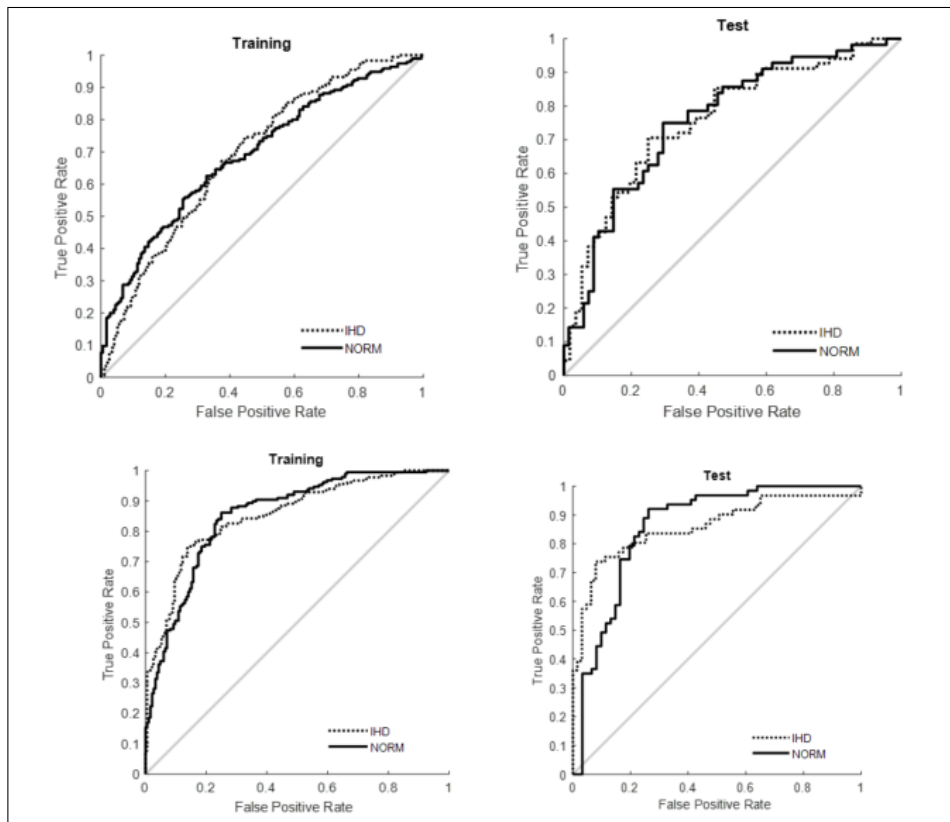

Figure 3.27: ROC curves in the training and test phases for the ANN2 (top panels) and ANN12 (bottom panels) producing the best accuracies.

Discussion

In the first study, with the validation test, the ANN modelling was more specific than sensitive tending to misclassify more IHD patients as normal. During the training, no normal subjects were classified as IHD while, during the validation test, about 10% of the normal subjects were identified as suffering from IHD. However, the difference between the percentage of false positive and of false negative, negligible during the training, was instead large during the validation test, implying that the ANN modelling was not able to balance sensitivity and specificity. This could be due to the low number of IHD subjects during the validation phase. In the second study, among the possible ANNs in which the first five components obtained using PCA were considered, the best network presented an accuracy of 80% and an AUC of 83% in the validation test. Although with a slightly lower performance, our network confirmed and extended the results of Dua et al. [96]. The difference between the accuracies could be due to the very small number of subjects considered in the previous work (only 20 subjects) in comparison with the large set (965 subjects) examined in this work. Moreover, in this paper only independent measurements were considered instead of repeated measurements on the same subjects, as in [96]. The highest ACC (82%) and AUC (85%) values were reached by the ANN2 scheme with only five parameters as input. In the last study, a greater variability in IHD than in normal subjects were shown with significantly greater values of RMSSD, HF, HFn and SD1 and lower values of LFn and LF/HF, partially in contrast with the results reported in the literature [72, 73] obtained by considering only normal beats. The presence of ectopic beats could explain this large variability as well as the presence of slightly higher (n.s.) values of LF in IHD since these beats, in RR series, often manifest as a derivative spike, leading to a flat power spectral density contribution at all frequencies (except the zero frequency). Moreover, since LF band is narrower than HF band, the presence of ectopic beats will increase HF power much more than LF power. Such a greater variability is confirmed by further new parameters examined in our work both in time domain, like greater SDNN, NN50 and pNN50 values, and in the non-linear domain like greater SD1/SD2, FD and lower Beta exponent values in IHD patients. The latter was calculated on the whole band (0.002-0.45Hz), contrarily to other authors [91] using only the very low frequency band (0.0001-0.03Hz) in which this parameter showed greater values in IHD patients. Our results also confirmed the outcomes of [83] presenting higher MeanRR values in IHD than in normal subjects. By using as inputs some linear, non-linear or a combination of these HRV parameters together with age and gender. The best ACC (71.8%) and AUC (71.5%) were obtained excluding among the linear parameters that were significantly correlated for more than 90% (ANN2). This accuracy was lower than that obtained by other authors [96, 108] applying ANN to HRV parameters on samples with a very limited number of subjects. Successively, adding as further input the LVEF feature, the classification performance of the examined ANNs increases of about 8% for

ACC and of 15% for AUC. The ANN presenting the highest ACC (79.8%) and AUC (86.0%) used as input five mixed linear and non-linear HRV parameters (MeanRR, LF, LF/HF, Beta exponent and SD2) together with age, gender and LVEF, highlighting the features that can be utilized for a high prognostic identification of IHD disease. This performance was comparable, as done for the first study, to the highest values reported in the literature [96] that developed an artificial neural network after the application of principal component analysis to only non-linear features extracted from HRV. In conclusion, these studies proved the power of ANN techniques for the classification of IHD patients using non-invasive parameters.

Main findings

The purpose of these studies were the identification of both ANN structures and HRV parameters, producing the best performance to identify IHD patients in a non-invasive way. The ANN applied to MeanRR, LF, LF/HF, Beta exponent, SD2 together with age and gender reached a maximum accuracy of 71.8% and, by adding as input LVEF, an accuracy of 79.8%.

3.3.2 CART for Dilated Cardiomyopathy [13SG] [14SG]

Materials and methods

Subjects

Initially, 773 (374 males, aged 63±19 and 399 females, aged 60±19) normal subjects and 199 (126 males, aged 57±14 and 73 females, aged 62±15) suffered from DCM were analyzed. Successively, a cohort of 1133 subjects, composed of 689 normal subjects (321 males, aged 62±15, and 368 females, aged 64±16), 263 patients affected by IHD (207 males, aged 71±10, and 56 females, aged 76±10) and 181 patients suffering from DCM (111 males, aged 59±12, and 70 females, aged 63±15) was analyzed.

HRV parameters

Several linear (in time and spectral) and non-linear parameters were evaluated on each subject and segment, containing either normal or ectopic beats. The mean value along the 24hrs was calculated for each parameter and the average among the subjects of each group was considered. In the second study, to describe the distribution of each parameter in the three groups, the mean (±SD) was computed for each group. Since the subject groups were independent and showed a non-Gaussian distribution, the significance of the difference between each pair of groups was evaluated by the Wilcoxon rank sum test with the Bonferroni correction considering p-values lower than 0.05.

Neural Network Classifier

In both studies, six different classifiers using three different subsets of normalized HRV parameters together with age and gender, adding or not the LVEF

parameter were considered. For the first combination of features the step-wise regression was applied to all the 17 parameters selecting those presenting p<0.05. In the second case, the variables correlated for less than 80% were selected while in the third case, all the 17 parameters were considered. For the other three trees (TREE 4, 5 and 6), we added to each of the first three combinations the LVEF parameter. In all studies, the training and test sizes were respectively 75% and 25% of the total number of data, randomly selected. The performance of each classifiers was calculated 1000 times, randomly changing the training and test datasets, the performance parameters were evaluated and the classifier producing the highest accuracy in the test phase was selected. In the second study the parameters were calculated separately for each group and weighted among the three groups as follow:

$$Weighted\,Sensitivity = \sum_{i=1}^{3}(\frac{TP_i}{(TP_i + FN_i)})p_i \tag{3.1}$$

$$Weighted\,Specificity = \sum_{i=1}^{3}(\frac{TN_i}{(TN_i + FP_i)})p_i \tag{3.2}$$

$$Weighted\,Precision = \sum_{i=1}^{3}(\frac{TP_i}{(TP_i + FP_i)})p_i \tag{3.3}$$

with

$$N = \sum_{i=1}^{3}(TP_i + TN_i + FP_i + FN_i) \tag{3.4}$$

$$p_i = (TP_i + FN_i)/N \tag{3.5}$$

i representing the $i-th$ group of subjects, TP_i the number of true positive, FP_i the number of false positive, FN_i the number of false negative, TN_i the number of true negative cases for the $i-th$ class. Moreover, the classification ACC was calculated as:

$$Accuracy = \sum_{i=1}^{3}(\frac{TP_i}{TN_i + FP_i + FN_i + TP_i})p_i \tag{3.6}$$

Finally, in the second study, it was evaluated the mean of the accuracy distribution for each TREE2 and the significance of the difference between each pair of the corresponding TREE2s with and without LVEF (i.e. TREE2.1 vs TREE2.4, TREE2.2 vs TREE2.5 and TREE2.3 vs TRE2.E6), the latter by using the Wilcoxon rank sum test because the subject groups were independent and showed a non-Gaussian distribution.

Tree Classifier	Input Features
TREE1	**pNN50**, **Beta exponent**, **gender**, age
TREE2	1°Comp, 2°Comp, 3°Comp, **4°Comp**, **5°Comp**
TREE3	MeanRR, SDNN, RMSSD, **NN50**, pNN50, LF, HF, LF/HF, LFn, HFn, **Beta exponent**, SD1, SD2, SD1/SD2, FD, gender, **age**
TREE4	**LF/HF**, gender, **LVEF**
TREE5	**1°Comp**, **2°Comp** , **3°Comp**, 4°Comp, 5°Comp
TREE6	MeanRR, SDNN, RMSSD, NN50, pNN50, LF, HF, LF/HF, LFn, HFn, **Beta exponent**, SD1, SD2, SD1/SD2, FD, gender, age, **LVEF**

Table 3.19: Input features utilized in the six cases in the first study. In bold the features selected by CART algorithm presented in the best sub-tree.

Results

About the first study, the features used by each best tree were reported in bold (Table 3.19) and the classification performances in the test phase of all the six classification trees were reported in Table 3.20. If the LVEF was added to the input, higher performances values (about 13% for accuracy) than those obtained without this clinical parameter, were achieved. The best accuracy, both in case with or without LVEF, was achieved by using parameters selected with stepwise regression (TREE1 and TREE4). The TREE4 produced the best performance with an accuracy of 97%. Figure shows the ROC curves of TREE1 and TREE4 in the test phases with AUC values of 67% and 95%, respectively.

	Test Set				
	SEN (%)	**SPE (%)**	**PRE (%)**	**ACC(%)**	**AUC (%)**
TREE1	18	99	89	84	67
TREE2	15	99	59	79	62
TREE4	90	99	95	97	95
TREE5	67	98	91	92	87
TREE6	90	98	93	96	94

Table 3.20: Classification performances on test set of the three models not including LVEF (TREE 1-3) and of the other three models considering LVEF (TREE4-6). In bold the highest values. SEN: sensitivity, SPE: specificity, PRE: precision, ACC: accuracy, AUC: area under the curve.
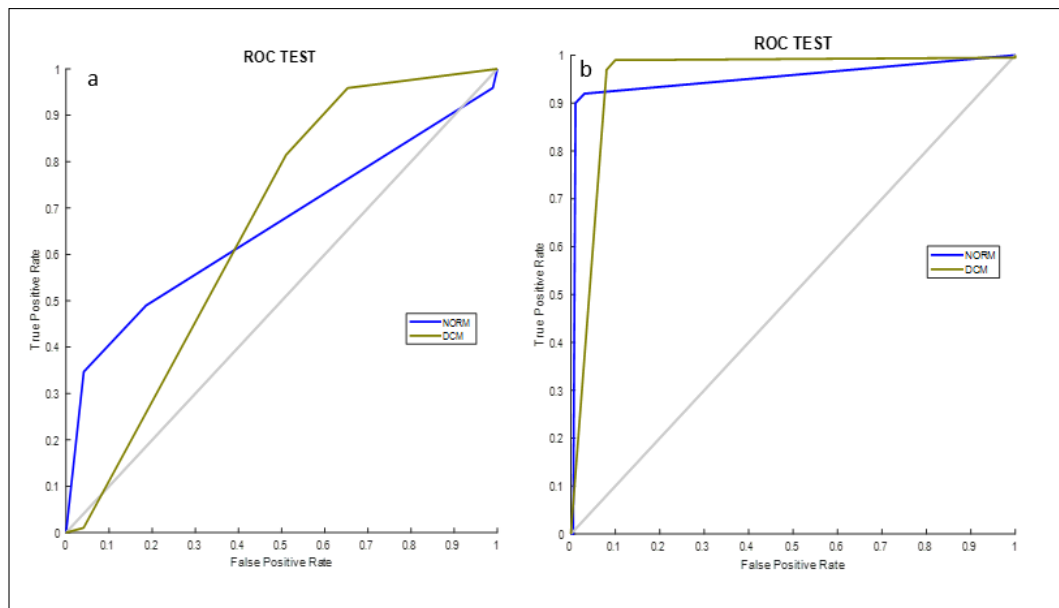
Figure 3.28: ROC test of TREE1 and TREE4.

In the second study, it was highlighted that almost all parameters showed significant differences between normal and the diseased groups, except for LF, and there were not significant differences between DCM and IHD values, except for MeanRR (Table 3.21). In particular, patients suffering from DCM and IHD presented higher values of the time domain measures than normal subjects. Among spectral parameters, LFn and LF/HF showed statistically higher values in normal subjects while HF and HFn presented an opposite behaviour. Considering the non-linear parameters, Beta exponent was statistically higher in normal than in DCM and IHD subjects while the opposite was showed by the Poincaré measurements and FD (Table 3.21).

Parameter	NORM subjects	DCM patients	IHD patients	p-value		
	Mean (± SD)	Mean (± SD)	Mean (± SD)	NORM vs DCM	NORM vs IHD	IHD vs DCM
MeanRR (ms)	877±138	880±130	942±145	1.1	<0.0001	<0.001
SDNN (ms)	71±53	85±61	87±68	<0.001	<0.05	1.6
RMSSD (ms)	38±91	61±111	71±115	<0.0001	<0.0001	2.8
NN50	50±73	65±85	70±90	<0.05	0.07	0.9
pNN50	0.15±0.21	0.19±0.23	0.21±0.27	<0.05	<0.05	1.7
LF (ms2)	460±2200	521±1500	350±2400	1.9	0.2	0.4
HF (ms2)	276±5800	626±6005	640±8100	<0.0001	<0.0001	2.8
LF/HF	2.06±1.90	1.22±1.30	0.97±1.05	<0.0001	<0.0001	0.2
LFn	0.59±0.22	0.44±0.21	0.40±0.20	<0.0001	<0.0001	0.3
HFn	0.44±0.22	0.58±0.21	0.60±0.20	<0.0001	<0.0001	0.3
Beta exp (ms2/Hz)	1.06±0.57	0.75±0.58	0.67±0.55	<0.0001	<0.0001	0.4
SD1 (ms)	34.80±34.10	45.50±40.90	46.40±42.40	<0.0001	<0.001	1.7
SD2 (ms)	83.10±54.20	92.80±57.30	95.50±70.10	<0.05	0.5	1.04
SD1/SD2	0.37±0.14	0.44±0.16	0.44±0.15	<0.0001	<0.0001	2.9
FD	1.53±0.16	1.63±0.16	1.63±0.15	<0.0001	<0.0001	0.5

Table 3.21: Mean (± SD) values of the HRV parameters and p-values of the difference between each pair of groups.

Table 3.22 presents the variables used as input for the six considered trees and highlighted in bold the features selected by the CART algorithm to generate the trees presenting the highest accuracy. When the LVEF was added to the input, higher performances values (about 10% for the maximum accuracy) than those obtained without this parameter, were achieved. A significant difference (p<0.00001) between the mean values of the accuracies of each pair of corresponding TREEs, with and without LVEF, was found. Table 3.23 shows the classification performances, together with the AUC values in the test phase, calculated both separately for each group and as weighted mean, of the two classification trees showing the best accuracy (i.e. TREE2.2 and TREE2.4). Finally, Figure 3.29 and Figure 3.30 showed the classification trees and the ROC curves of the three with the best accuracy.

	Parameters
TREE2.1	**MeanRR**,**SDNN**,**LFn**, **FD**, gender,age
TREE2.2	**MeanRR**, **SDNN**, **LF**,**LF/HF**,**Beta exp**, gender, **age**
TREE2.3	MeanRR, SDNN, RMSSD, **NN50**, pNN50, **LF**, **HF**, LF/HF, LFn, **HFn**, Beta exp, SD1, SD2, SD1/SD2, FD, **gender**, **age**
TREE2.4	MeanRR, **pNN50**, LF/HF, **FD**, **gender**, **age**, **LVEF**
TREE2.5	MeanRR, SDNN, LF, LF/HF, Beta exp,**gender**, **age**,**LVEF**
TREE2.6	MeanRR, **SDNN**, RMSSD, NN50, pNN50, **LF**, HF, LF/HF, LFn, **HFn**, Beta exp, SD1, SD2, **SD1/SD2**, **FD**, gender, **age**, **LVEF**

Table 3.22: Input features utilized in the six cases, in the second study. In bold the features selected by CART algorithm presented in each best sub-tree.

	TREE2.2	TREE2.4
SEN-NORM (%)	88.4	98.3
SEN-DCM (%)	11.1	71.1
SEN-IHD (%)	69.7	41.5
SPE-NORM (%)	56.8	66.4
SPE-DCM (%)	99.2	95.8
SPE-IHD (%)	86.2	96.8
PRE-NORM (%)	76.0	82.1
PRE-DCM (%)	71.4	76.2
PRE-IHD (%)	60.5	79.4
AUC-NORM (%)	78.0	86.0
AUC-DCM (%)	60.0	91.0
AUC-IHD (%)	86.0	80.0
Weighted SEN (%)	71.7	80.9
Weighted SPE(%)	70.4	78.1
Weighted PRE (%)	71.7	80.5
Weighted AUC (%)	77.0	86.0
Max ACC (%)	71.7	80.9
Median ACC (%)	61.5	74.2

Table 3.23: Classification performances, for each subject group and weighted among groups, for the best trees considering (TREE2.4) or not (TREE2.2) the LVEF parameter.
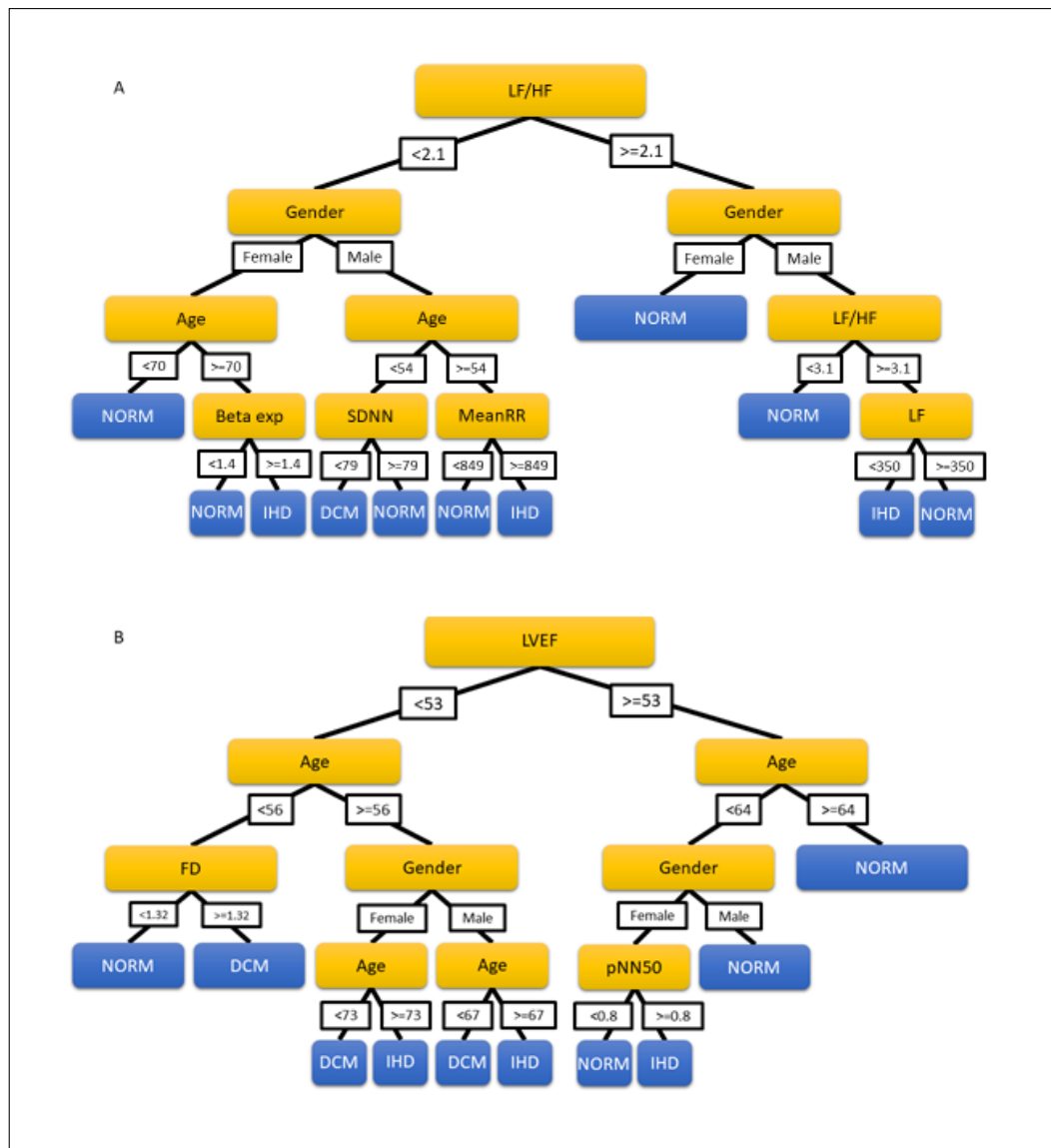
Figure 3.29: Classification trees presenting the best accuracy in cases of use or not of LVEF. A: TREE2.2, without LVEF, B: TREE2.4, with LVEF. Each node is the graphical representation of a set of "if...then" rules thus for example, in TREE2.2, if LF/HF is less than 2.1, if the gender is female and if age is less than 70 years old, the subject is classified as normal.
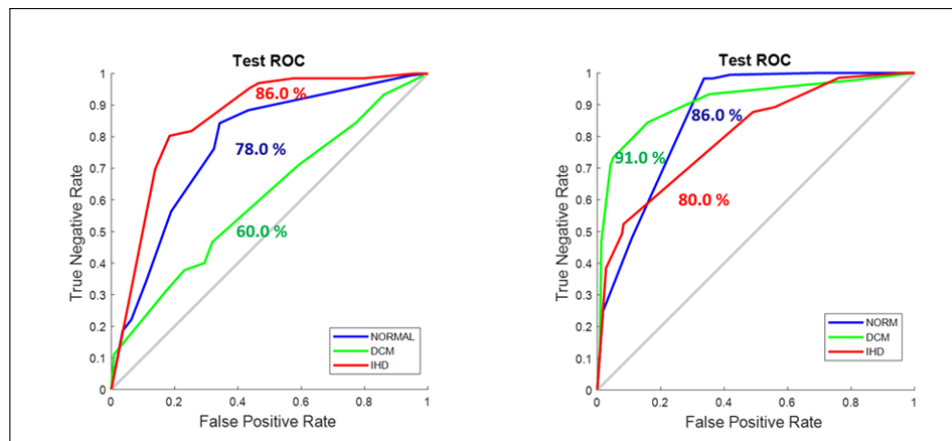
Figure 3.30: Receiver Operating Characteristic curves and corresponding %Area Under the Curve values for each group during the Test phase of TREE2.2 (left) and TREE2.4 (right).

Discussion

In the first study, the CART algorithm was used to identify DCM from normal subjects using as features a combination of HRV parameters, age, gender and the clinical outcome (LVEF). The best performance was achieved applying stepwise regression obtaining an accuracy of 97%. In the second study, considering normal subjects, IHD and DCM patients, almost all HRV parameters showed significant differences between normal subjects and patients groups, except for LF, while there were not significant differences between DCM and IHD values, except for MeanRR. This latter parameter presented statistically higher values in patients than in normal subjects, confirming the results of [86]. Considering IHD patients, higher values on time domain, HFn and Poincaré parameters and lower values of LF/HF and LFn parameters than normal subjects (Table 3.21) were opposite to the findings presented in [72] and in [73]. Moreover, regarding DCM patients, time domain and Poincaré parameters (Table 3.21) showed greater values than normal subjects with an opposite behavior compared to that reported by [73, 82, 83]. On the other hand, LFn and LF/HF spectral parameters decreased in DCM compared to normal subjects and HFn values increased, according to the results of [73, 83] (Table 3.21). Finally, our results showed significantly greater FD values and lower Beta exponent values in patients than in normal subjects, supporting the evidence of a greater total heart rate variability in patients than in normal subjects (Table 3.21). However, these differences could be due to the presence of both normal and ectopic beats considered in comparison with the analyses carried out in the literature in which only normal beats were taken into account [72, 73, 82, 83]. The greater number of anomalous beats in patients produces a greater total variability in the RR sequence with consequent higher values in HF than in LF components as well as in time domain, FD and Poincaré parameters values in patients than in normal.

Nowadays, different mathematical approaches for DSS [96, 109, 110] have been proposed for the automatic classification of heartbeats. Among them, only Dua et al [96] used the classification and regression tree analysis to identify IHD patients, analyzing a small cohort of subjects. Other authors [109, 110] identified heart beat or segments of DCM patients using HRV parameters as input of complex classifier in which the CART was combined with other machine learning techniques but taking data from a diagnostic ECG databases[1] without developing a DSS capable of classifying subjects.

In conclusion, these are the first studies able to classify cardiovascular diseases using CART method on a large dataset, point out that this machine learning technique can be a resource in optimizing early differential diagnosis configuring it as a useful non-invasive clinical and possibly screening tool.

Main findings

Initially, the aim of the first study was to verify if parameters derived from HRV, applied to a CART could be able to distinguish normal and idiopathic Dilated Cardiomyopathy patients. Linear and non-linear HRV parameters and some clinical parameters (gender, age and LVEF) evaluated on a large cohort of subjects were considered. By using principal component analysis and stepwise fit regression, the original parameters were reduced and used as inputs of CART obtaining an ACC of 97%. In the second study, the CART algorithm was applied to distinguish DCM and IHD from normal subjects and between them. Analysing the HRV parameters, most of them showed significant differences between normal and diseased groups. There were not significant differences between HRV values in DCM and IHD patients, except for the MeanRR. The features selections were applied, and applying CART algorithms to several combinations of HRV parameters with age and gender, an accuracy of 71.7% was reached. Moreover, adding LVEF as inputs, the algorithm was able to distinguish the three subgroups with an accuracy of 80.9%.

---

[1]www.physionet.org

Conclusion

The aim of my research activity was the application of biomedical signal processing as an instrument able to support clinicians in different settings such as: to correctly evaluate physiological states over 24hrs, to quantify the cardiovascular variations over 24hrs due to the influence of risk factors and to early identify subjects with different pathologies from the analysis of cardiovascular signal over 24hrs. The activity could be subdivided in three sections concerning three different topics.

In particular in the first section the HR and BP circadian rhythms and their relationship were examined in detail, applying linear regression analysis in characteristic periods of time over 24hrs, considering normotensive and hypertensive subjects. In the post-prandial period, the decreasing HR and increasing BP values suggested that the mechanisms that regulate the circadian behaviour of the two signals are independent. To better describe the circadian trend of the relationship between BP and HR in both subject groups, a regression line for each hour over 24hrs was applied. The large variability of slopes and intercepts found over 24hrs suggested that the conventional punctual measurements, such as office or the simple average over 24hrs, are not sufficient to accurately describe this cardiovascular relationship.

In the second section, the influence of some risk factors, such as age, gender, smoking, obesity and dyslipidaemia on cardiovascular signals was studied. Initially, the influence of age on RR variability over 24hrs was analysed in subjects from 15 to 90 years old, founding a similar circadian rhythm in all linear and non-linear parameters with a significant differences between day and night. Moreover, a parabolic relationship between these parameters and age, not yet reported in the literature, was found with an opposite trend in subjects over about 60 years old. Finally, analysing the gender effect on the relationship between BP/HR acting on both the sympathetic/parasympathetic systems, the results underlined a different trend in the BP/HR relationship between gender only during the night time. The effects of smoking on BP and HR were examined using linear regression analysis highlighting that the slopes could be used to quantify both the morning surge and the night fall one, associated with cardiovascular effects, in both subject groups. Moreover, the variations due to the presence of each risk factor on HR circadian rhythm highlighted that each factor influences the rhythm differently both in normotensive and hypertensive

subjects. Since the circadian HR trend presented different characteristic periods, a linear regression analysis was applied on each period underlining that during awakening there were no differences between subjects with and without risk factors from 10:00 to 04:00 and that smoking increased HR values while dyslipidemia decreased HR values. In addition, since BP is not only inevitably affected by ageing but also by the presence of other risk factors, the relationship between BP/age changes with the presence of risk factors in hypertensive and normotensive subjects was studied. The regression line fitted this relationship, considering values taken from office and ABPM measurements and the results highlighted higher slopes in the SBP/age relation in hypertensive than in normotensive subjects in both measurements.

In the last section, in order to identify cardiovascular diseases, different mathematical approaches for developing DSS based on matching learning techniques were examined. Initially, artificial neural networks (ANNs), using as inputs all the linear and non-linear HRV parameters evaluated along 24hrs together with age and gender, were developed and tested on a large sample of subjects. Successively, the clinical parameter named LVEF was added as an input and, as a consequence, the classification performances improved. In order to reduce the number of inputs, a feature selection (preserving the variance) was done by using stepwise regression, principal component analysis and correlation coefficient. The ANN with the highest accuracy (79.8%) to distinguish normal from Ischemich Heart patients, presented as inputs MeanRR, LF, LF/HF, Beta exponent, SD2 together with age, gender and LVEF. Successively, another machine learning technique such as the CART algorithm was applied to identify Dilated Cardiomyopathy patients from normal ones by using the inputs previous described, achieving an accuracy of 97%. Finally, the HRV parameters over 24hrs were also evaluated in IHD, pointing out that most of them showed significant differences between normal and diseased groups with not significant differences between HRV values in DCM and IHD patients, except for MeanRR. Applying CART algorithms to identify subjects belonging to one of the three groups, an accuracy of 71.7% was reached. Moreover, adding LVEF as inputs, the algorithm was able to distinguish the three subgroups with an accuracy of 80.9%. All these studies provide a deep insight into how a combination of non-invasive parameters, obtained from signal processing and from echocardiography (LVEF) together with age and gender, could be exploited to reliably detect the presence of subjects affected by cardiovascular diseases.

Future research

Future research concerning the analysis of cardiovascular rhythms could compare the different technologies that ensure continuous non-invasive monitoring, in order to evaluate the limits of each technology in specific clinical areas. In particular, the differences between the two most common measures, such as office and ABPM, could be carried out on the two periods of time over 24hrs (day and night) commonly examined in the literature as well as on the characteristic periods highlighted in this thesis.

Furthermore, from a clinical point of view, blood pressure is an epidemic factor, that is, a determining factor in order to create an epidemic condition of the disease. For this reason, it can be conducted a retrospective study in order to determine, in detail, the influence of arterial hypertension and other risk factors on the onset of arterial plaques. In regards to the application of machine learning techniques able to identify cardiovascular disease, further efforts will have to be made to improve classification performances of those diagnostic algorithms. To do that, emerging non-invasive clinical parameters such as global longitudinal strain could be integrated in order to obtain an ever earlier diagnosis and improve patients prognosis. In addition to this, other machine learning techniques could be exploited and implemented such as support vector machine and Kth nearest-neighbours.

Acknowledgement

I would like to express my gratitude to prof. Accardo for his guidance during my years as a graduate and post graduate student. I am grateful to Dr Lorenzo Pascazio of the Geriatric Department of ASUGI for sharing his medical knowledge with me. I acknowledge all of the specialists of the Cardiovascular Department of ASUGI for their pleasant and fruitful collaboration. I owe special gratitude to my colleagues for all they have contributed so unselfishly. In particular, I wish to thank my PhD colleague, Aleksandar, for the good moment we have shared during these years. I would also like to thank my senior researcher Milos for having transmitted to me the passion for the research. I am grateful to all my "triestini" friends, flatmates and old friends for bearing with me through the last years and for the invaluable encouragement provided during this work. I am sure that without them this work would not have been possible. Finally, I wish to thank my parents, Alvaro and Deanna, for supporting and loving me. To them I dedicate this thesis.

My PhD Bibliography

This thesis presents results published in:

1SG Silveri G., Pascazio L., Accardo A., Relationship between blood pressure and heart rate circadian rhythms in normotensive and hypertensive subjects, *Computing in Cardiology*, pp. 1-4, 2018

2SG Silveri G., Pascazio L., Sabbadini G., Guerra M. and Accardo A., Differences in Circadian Rhythms of Blood Pressure and Heart Rate Among Hypertensive and Normal Blood Pressure Subjects, *15th Mediterranean Conference on Medical and Biological Engineering and Computing*, 76, pp. 142-149, 2019

3SG Silveri G., Pascazio L., Accardo A., Influence of the time of day on the relationship between heart rate and blood pressure, *J Med Eng Technol*, in publication

4SG Silveri G., Pascazio L., Accardo A., Effects of Smoking on HR Circadian Rhythm in Hypertensive and Non Hypertensive Subjects, *11th Conference of the European Study Group on Cardiovascular Oscillations*, 9158053, 2020

5SG Silveri G., Pascazio L., Miladinović A., Ajčević M., Accardo A., Smoking effect on the circadian rhythm of blood pressure in hypertensive subjects, *7th Congress of the National Group of Bioengineering*, in publication

6SG Silveri G., Pascazio L., Ajčević M., Accardo A., Influence of smoking and other cardiovascular risk factors on HR circadian rhythm in normotensive and hypertensive subjects, *PloseOne*, in publication

7SG Silveri G., Pascazio L., Ajčević M., Miladinović A., Accardo A., Influence of some cardiovascular risk factors on the relationship between age and blood pressure, *18th Nordic-Baltic Conference on Biomedical Engineering and Medical Physics*, 2020

8SG Accardo A., Merlo M., Silveri G., Del Popolo L., Della Libera L., Restivo L., Cinquetti M., Cannatà A., Sinagra G., Influence of ageing on circadian rhythm of HRV in healthy subjects, *J Cardiovasc Med*, 26 Aug, 2020

9SG Silveri G., Pascazio L., Ajčević M., Miladinović A., Accardo A., Influence of the gender on the relationship between heart rate and blood pressure, *8th European Medical and Biological Engineering Conference*, 2020

10SG Silveri G., Merlo M., Restivo L., De Paola B., Miladinovic A., Ajcevic M., Sinagra G., Accardo A., Identification of Ischemic Heart Disease by using machine learning technique based on parameters measuring Heart Rate Variability *28th European Signal Processing Conference*, pp. 1310-1312, 2020

11SG Silveri G., Merlo M., Restivo L., Sinagra G., Accardo A., Novel Classification of Ischemic Heart Disease Using Artificial Neural Network, *Computing in Cardiology*, 2020

12SG Accardo A., Silveri G., Merlo M., Restivo L., Ajčević M., Sinagra G., Detection of subjects with Ischemic Heart Disease by using machine learning technique based on Heart Rate Total Variability parameters, *Physiological Measurement*, 20 Octob, 2020

13SG Silveri G., Merlo M., Restivo L., Ajčević M., Sinagra G., Accardo A., A big – data classification tree for decision support system in the detection of dilated cardiomyopathy using heart rate variability, *Procedia Computer Science*, 176, pp. 2940-2048, 2020

14SG Accardo A., Restivo L., Silveri G., Merlo M., Sinagra G, Identification of Dilated Cardiomyopathy and Ischemic Heart Disease by machine learning system applied to Heart Rate Total Variability, *JACC*, in publication

In the last years I worked also on other topics in the field of biomedical signal processing. The results of these studies are reported in the following publications not included in this thesis:

15SG Silveri G., De Dea F., Accardo A., Influence of dysgraphia on kinematic characteristics of handwriting in Italian primary school children, *World Congress on Medical Physics and Biomedical Engineering*, 68, pp. 241-245, 2018

16SG Silveri G., Accardo A., Handwriting parametrization for dysgraphia identification, *6th Congress of the National Group of Bioengineering*,2018

17SG Marino S.,Silveri G., Miladinović A., Ajčević M., Accardo A., Linear and Non-linear Analysis of EEG During Sleep Deprivation in Subjects with and Without Epilepsy, *15th Mediterranean Conference on Medical and Biological Engineering and Computing*, 76, pp. 125-132, 2019

18SG Miladinović A., Ajčević M., Battaglini P.P., Silveri G., Jarmolawska J., Accardo A., Slow Cortical Potential BCI Classification Using Sparse

Variational Bayesian Logistic Regression with Automatic Relevance Determination, *15th Mediterranean Conference on Medical and Biological Engineering and Computing*, 76, pp. 1853-1860, 2019

19SG   Silveri G., Accardo A., Handwriting Kinematic Differences Between Copying and Dictation, *15th Mediterranean Conference on Medical and Biological Engineering and Computing*, 76, pp. 142-149, 2019

20SG   Ajčević M., Miladinović A., Silveri G.,Buoite Stella A., Caruso P., Ukmar M., Naccarato M., Cuzzocrea A.M., Manganotti P., Accardo A., A Big-Data Variational Bayesian Framework for Supporting the Prediction of Functional Outcomes in Wake-Up Stroke Patients, *Lecture Notes in Computer Science*, 2020

21SG   Miladinović A., Ajčević M., Jarmolawska J., Silveri G., Battaglini P.P., Accardo A., Performance of EEG Motor-Imagery based spatial filtering methods: A BCI study on Stroke patients, *Procedia Computer Science*, 176, pp. 2840 - 2848, 2020

22SG   Miladinović A., Ajčević M., Busan P., Jarmolawska J., Silveri G., Mezzarobba S.,Battaglini P.P., Accardo A., Transfer Learning improves MI BCI models classification accuracy in Parkinson's disease patients, *28th European Signal Processing Conference,2020*

23SG   Miladinović A., Ajčević M., Silveri G., Accardo A., Performance of Dual-Augmented Lagrangian Method and Common Spatial Patterns applied in classification of Motor-Imagery BCI, *7th Congress of the National Group of Bioengineering*, in publication

25SG   Ajčević M., Miladinović A., Naccarato M., Silveri G., Caruso P., Accardo A., Manganotti P., Wireless EEG in hyper-acute ischemic stroke: Correlation between neurophysiological alterations and CTP total hypoperfused volume, *Procedia Computer Science*, 176, pp. 2923-2929, 2020

26SG   Ajčević M., Miladinović A., Stragapede L., Silveri G., Caruso P., Naccarato M., Manganotti P., Accardo A., Correlation between hyper-acute EEG alterations and 7-Day NIHSS score in thrombolysis treated ischemic stroke patients, *7th Congress of the National Group of Bioengineering*, in publication

27SG   Miladinović A., Ajčević M., Jarmolawska J., Colussi M., Silveri G., Battaglini P.P., Accardo A., Effect of power feature covariance shift on BCI spatial-filtering techniques: A comparative study, *Computer methods and programs in biomedicine*, 198,15808,2020

28SG   Accardo A., Pascazio L., Constantinides F., Gorza F., Silveri G., Influ-
       ence of hypertension and other risk factors on the onset of sublingual
       varices, *BMC Oral Helth*, in publication

Bibliography

[1] Aschoff J. Circadian rhythms in man. *Science*, 148(3676):1427–1432, month 1965.

[2] deBoer R.W. et al. Hemodynamic fluctuations and baroreflex sensitivity in humans: a beat-to-beat model. *Am J Physiol*, 253(3):H680–9, 1987.

[3] Johansen C.D. et al. Resting, night-time, and 24h heart rate as markers of cardiovascular risk in middle-aged and elderly men and women with no apparent heart disease. *Eur Heart J*, 34(23):1732–9, 2013.

[4] Guyenet P.G. The sympathetic control of blood pressure. *Nat Rev Neurosci*, 7(5):335–46, 2006.

[5] Palatini P. et al. Heart rate and the cardiovascular risk. *J Hypertens*, 15(1):3–17, 1997.

[6] Diaz K.M. et al. Visit-to-visit variability of blood pressure and cardiovascular disease and all-cause mortality: a systematic review and meta-analysis. *Hypertension*, 64(5):965–82, 2014.

[7] Williams B et al. 2018 ESC/ESH guidelines for the management of arterial hypertension the task force for the management of arterial hypertension of the European Society of Cardiology (ESC) and the European Society of Hypertension (ESH). *Eur Heart J*, 38(number):3021–3104, 2018.

[8] Malik M. at al. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. *Eur Heart J*, 17(3):354–381, 1996.

[9] Kannel W.B. et al. Heart rate and cardiovascular mortality: the framingham study. *Am Heart J*, 113(6):1486–94, 1987.

[10] Gillman M.W. et al. Influence of heart rate on mortality among persons with hypertension: the framingham study. *Am Heart J*, 125(4):1148–54, 1993.

[11] Pickering T.G. et al. Ambulatory blood-pressure monitoring. *N Engl J Med*, 1(354):2368–74, 2006.

[12] Mancia G. Short- and long-term blood pressure variability: present and future. *Hypertension*, 60(2):512–7, 2012.

[13] Palatini P. et al. Added predictive value of night-time blood pressure variability for cardiovascular events and mortality. The ambulatory blood pressure–international study. *Hypertension*, 64(3):487–493, 2014.

[14] Hansen T.W. et al. Prognostic value of ambulatory heart rate revisited in 6928 subjects from 6 populations. *Hypertension*, 52(2):229–235, 2008.

[15] Mancia G. et al. Ambulatory blood pressure monitoring and organ damage. *Hypertension*, 36(5):894–900, 2000.

[16] Millar-Craig M.W. et al. Circadian variation of blood-pressure. *Lancet*, 15(1):795–7.

[17] Mancia G. et al. Blood pressure and heart rate variabilities in normotensive and hypertensive human beings. *Circ Res*, 53(1):96–104, 1983.

[18] Richards A.M. et al. Diurnal patterns of blood pressure, heart rate and vasoactive hormones in normal man. *Clin Exp Hypertens A*, 8(2):153–66, 2009.

[19] Littler W.A. et al. The variability of arterial pressure. *Am Heart J*, 95(2):180–186, 1978.

[20] Zakopoulos N. et al. Effect of hypotensive drugs on the circadian blood pressure pattern in essential hypertension: a comparative study. *Cardiovasc Drugs Ther*, 11(number):795–799, 1997.

[21] Verdecchia P. Prognostic value of ambulatory blood pressure : current evidence and clinical implications. *Hypertension*, 35(3):844–51, 2000.

[22] Madden J.M. et al. Exploring diurnal variation using piecewise linear splines: an example using blood pressure. *Emerg Themes Epidemiol*, 14(1):pages, 2017.

[23] Morcet J.F. at al. Associations between heart rate and other risk factors in a large french population. *J Hypertens*, 17(12):1671–6, 1999.

[24] Zhang J. et al. Anthropometric, lifestyle and metabolic determinants of resting heart rate.a population study. *Eur Heart J*, 20(2):103–10, 1999.

[25] Pickering T.G. et al. Ambulatory blood-pressure monitoring. *N Engl J Med*, 354(22):2368–74, 2006.

[26] Goldberger J.J. et al. Relationship of blood pressure to heart rate in isolated systolic hypertension. *J Investig Med*, 59(8):1228–32, 2011.

[27] Koroboki E. et al. Circadian variation of blood pressure and heart rate in normotensives, white-coat, masked, treated and untreated hypertensives. *Hellenic J Cardiol*, 53(6):423–8, 2012.

[28] Virdis A. et al. Cigarette smoking and hypertension. *Curr Pharm Des*, 16(23):2518–25, 2010.

[29] Ambrose J.A. The pathophysiology of cigarette smoking and cardiovascular disease: An update. *JACC*, 43(10):1731–37, 2004.

[30] Cryer P.E. et al. Norepinephrine and epinephrine release and adrenergic mediation of smoking-associated hemodynamic and metabolic events. *N Engl J Med*, 295(11):573–7, 1976.

[31] Tsuda A. et al. Cigarette smoking and psychophysiological stress responsiveness: effects of recent smoking and temporary abstinence. *Psychopharmacology*, 126:226–233, 1996.

[32] Al-Safi S.A. Does smoking affect blood pressure and heart rate? *Eur J Cardiovasc Nurs*, 4(4):286–9, 2005.

[33] Bolinder G. et al. Ambulatory 24-h blood pressure monitoring in healthy, middle-aged smokeless tobacco users, smokers, and nontobacco users. *Am J Hypertens*, 11(10):1153–63, 1998.

[34] Tuomilehto J. et al. Decreased coronary heart disease in hypertensive smokers. Mortality results from the maphy study. *Hypertension*, 13(6):773–80, 1989.

[35] Bang L.E. et al. Do we undertreat hypertensive smokers? A comparison between smoking and non-smoking hypertensives. *Blood Press Monit*, 5(5):271–4, 2000.

[36] Verdecchia P. et al. Cigarette smoking, ambulatory blood pressure and cardiac hypertrophy in essential hypertension. *J Hypertens*, 13(10):1209–15, 1995.

[37] Sørensen K. et al. Increased systolic ambulatory blood pressure and microalbuminuria in treated and non-treated hypertensive smokers. *Blood Press*, 13(6):362–8, 2004.

[38] Lee J.F. et al. Elevated resting heart rate and reduced orthostatic tolerance in obese humans. *Clin Auton Res*, 24(1):39–46, 2014.

[39] Perlini S. et al. Pulse pressure and heart rate in patients with metabolic syndrome across europe: insights from the good survey. *J Hum Hypertens*, 27(7):412–6, 2013.

[40] Rossi R.C. et al. Impact of obesity on autonomic modulation, heart rate and blood pressure in obese young people. *Auton Neurosci*, 193:138–41, 2015.

[41] Grassi G. et al. Sympathetic activation in obese normotensive subjects. *Hypertension*, 25(4):560–3, 1995.

[42] Freitas Júnior I.F. et al. Resting heart rate as a predictor of metabolic dysfunctions in obese children and adolescents. *BMC Pediatr*, 12(12):5, 2012.

[43] Kotsis V. et al. Impact of obesity on 24-hour ambulatory blood pressure and hypertension. *Hypertension*, 45(4):602–7, 2005.

[44] Sun J.C. et al. Elevated resting heart rate is associated with dyslipidemia in middle-aged and elderly chinese. *Biomed Environ Sci*, 27(8):601–5, 2014.

[45] Lee J.S. et al. Triglyceride and hdl-c dyslipidemia and risks of coronary heart disease and ischemic stroke by glycemic dysregulation status: The strong heart study. *Diabetes Care*, 40(4):529–537, 2017.

[46] Palatini P. et al. Impact of increased heart rate on clinical outcomes in hypertension: implications for antihypertensive drug therapy. *Drugs*, 66(2):133–44, 2006.

[47] Gribbin B. et al. Effect of age and high blood pressure on baroreflex sensitivity in man. *Circ Res*, 29(4):424–31, 1971.

[48] Franklin S.S. et al. Hemodynamic patterns of age-related changes in blood pressure the framingham heart study. *Circulation*, 96(1):308–315, 1997.

[49] Mancia G. et al. Ambulatory blood pressure normality: results from the pamela study. *J Hypertens*, 13(12):1377–90, 1995.

[50] Staessen J.A. et al. Short report: ambulatory blood pressure in normotensive compared with hypertensive subjects.the ad-hoc working group. *J Hypertens*, 11(11):1289–97, 1993.

[51] Salice P. et al. Differences between office and ambulatory blood pressures in children and adolescents attending a hospital hypertension clinic. *J Hypertens*, 31(11):2165–75, 2013.

[52] Ishikawa J. et al. Age and the difference between awake ambulatory blood pressure and office blood pressure: a meta-analysis. *Blood Press Monit*, 16(4):159–67, 2011.

[53] Conen D. et al. Age-specific differences between conventional and ambulatory daytime blood pressure values. *Hypertension*, 64(5):1073–9, 2014.

[54] Banegas J.R. et al. Clinic versus daytime ambulatory blood pressure difference in hypertensive patients: the impact of age and clinic blood pressure. *Hypertension*, 69(2):211–219, 2017.

[55] Mancia G. et al. Relationship of office, home, and ambulatory blood pressure to blood glucose and lipid variables in the pamela population. *Hypertension*, 45(6):1072–7, 2005.

[56] Primatesta P. et al. Association between smoking and blood pressure: evidence from the health survey for england. *Hypertension*, 37(2):187–93, 2001.

[57] Ettehad D. et al. Blood pressure lowering for prevention of cardiovascular disease and death: a systematic review and meta-analysis. *Lancet*, 5(387):957–967, 2016.

[58] Dhingra R at al. Age as a risk factor. *Med Clin North Am*, 2012.

[59] Nicolini P. et al. The prognostic value of heart rate variability in the elderly, changing the perspective: from sympathovagal balance to chaos theory. *Pacing Clin Electrophysiol*, 35(5):622–38, 2012.

[60] Bonnemeier H.et al. Circadian profile of cardiac autonomic nervous modulation in healthy subjects: differing effects of aging and gender on heart rate variability. *J Cardiovasc Electrophysiol*, 14(8):791–9, month 2003.

[61] Hartikainen J.E. et al. Distinction between arrhythmic and nonarrhythmic death after acute myocardial infarction based on heart rate variability, signal-averaged electrocardiogram, ventricular arrhythmias and left ventricular ejection fraction. *J Am Coll Cardiol*, 28(2):296–304, 1996.

[62] Fauchier L. et al. Heart rate variability in idiopathic dilated cardiomyopathy: characteristics and prognostic value. *J Am Coll Cardiol*, 30(4):1009–14, 1997.

[63] Monfredi O. et al. Complexities in cardiovascular rhythmicity: perspectives on circadian normality, ageing and disease. *Cardiovasc Res*, 115(11):1576–1595, month 2019.

[64] Abdel-Rahman A.R. et al. Gender-related differences in the baroreceptor reflex control of heart rate in normotensive humans. *J Appl Physiol*, 77(2):606–13, 1994.

[65] Fricker J. Baroreflex sensitivity lower in women than men. *Lancet*, 348:257, 1996.

[66] Sa Cunha R. at al. Association between high heart rate and high arterial rigidity in normotensive and hypertensive subjects. *J Hypertens*, 15(12):1423–30, 1997.

[67] Khoury S. at al. Ambulatory blood pressure monitoring in a nonacademic setting. effects of age and sex. *Am J Hypertens*, 5(9):616–23, 1992.

[68] Thayer J.F. et al. Gender differences in the relationship between resting heart rate variability and 24-hour blood pressure variability. *Blood Press*, 25(1), 2016.

[69] Hermida R.C. at al. Modeling the circadian variability of ambulatorily monitored blood pressure by multiple-component analysis. *Chronobiol Int*, 19(2):461–81, 2002.

[70] Jaquet F. et al. Effects of age and gender on ambulatory blood pressure and heart rate. *J Hum Hypertens*, 12(4):253–7, 1998.

[71] Kleiger R.E. et al. Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. *Am J Cardiol*, 1(59):256–62, 1987.

[72] Acharya U.R. et al. Linear and nonlinear analysis of normal and cad-affected heart rate signals. *Comput Methods Programs Biomed*, 113(1):55–68, 2014.

[73] Voss A. et al. Methods derived from nonlinear dynamics for analysing heart rate variability. *Philos Trans A Math Phys Eng Sci*, 28(367):277–96, 2009.

[74] Jameson J.L. et al. *Harrison's Principles of Internal Medicine,20e*. McGraw-Hill Education, 2015.

[75] Goldenberg I. et al. Heart rate variability for risk assessment of myocardial ischemia in patients without known coronary artery disease: The HRV-DETECT (heart rate variability for the detection of myocardial ischemia) study. *J Am Heart Assoc*, 17(8):1–10, 2019.

[76] Knuuti J. et al. Esc scientific document group. 2019 esc guidelines for the diagnosis and management of chronic coronary syndromes. *Eur Heart J*, 14(41):407–477, 2020.

[77] Elliott P. Diagnosis and management of dilated cardiomyopathy. *Heart*, 84:106, 2000.

[78] Bozkurt B. et al. Current diagnostic and treatment strategies for specific dilated cardiomyopathies: a scientific statement from the american heart association. *Circulation*, 6(134), 2016.

[79] Rajendra Acharya U. et al. Heart rate variability: a review. *Med Biol Eng Comput*, 44(12):1031–51, 2006.

[80] Guzzetti S. et al. Heart rate variability in chronic heart failure. *Auton Neurosci*, 20(90):102–5, 2001.

[81] Hadase M. et al. Very low frequency power of heart rate variability is a powerful predictor of clinical prognosis in patients with congestive heart failure. *Circulation Journal*, 68(4):343–347.

[82] Yi G. et al. Heart rate variability in idiopathic dilated cardiomyopathy: relation to disease severity and prognosis. *Heart*, 77(2):108–14, 1997.

[83] Carvajal R. et al. Correlation dimension analysis of heart rate variability in patients with dilated cardiomyopathy. *Comput Methods Programs Biomed*, 78(2):133–40, 2005.

[84] Wessel N. et al. Long-term symbolic dynamics for heart rate variability analysis in patients with dilated cardiomyopathy. *Computers in Cardiology*, 26:253–256, 1999.

[85] Malberg H. et al. Short-term heart rate turbulence analysis versus variability and baroreceptor sensitivity in patients with dilated cardiomyopathy. *Indian Pacing Electrophysiol J*, 4(4):162–175, 2004.

[86] Voss A. et al. Methods derived from nonlinear dynamics for analysing heart rate variability. *Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences*, 367(1887):277–96, 2008.

[87] Adetiba E. et al. Automated detection of heart defects in athletes based on electrocardiography and artificial neural network. *Cogent Engineering*, 4(1), 2017.

[88] Ali M. et al. Detection of cardiomyopathy using multilayered perceptron network. *2012 IEEE 8th International Colloquium on Signal Processing and its Applications*, 2012.

[89] Ghosh D. et al. Wavelet aided svm analysis of ecg signals for cardiac abnormality detection. In *INDICON*, 2006.

[90] Ahmad Shukri M.H. et al. Investigation on elman neural network for detection of cardiomyopathy. In *IEEE Control and System Graduate Research Colloquium*, 2012.

[91] Guzzetti S. et al. Linear and non-linear 24 h heart rate variability in chronic heart failure. *Auton Neurosci*, 28(86):114–9, 2000.

[92] Rajendra Acharya U. et al. Heart rate variability: a review. *Med Biol Eng Comput*, 44(12):1031–51, 2006.

[93] Christov I. et al. Comparative study of morphological and time-frequency ecg descriptors for heartbeat classification. *Medical Engineering and Physics*, 28(9):876–887, 2006.

[94] Rodríguez-Sotelo J.L. et al. *Recognition of Cardiac Arrhythmia by Means of Beat Clustering on ECG-Holter Recordings*. Intechopen, 2012.

[95] Llamedo M. et al. Heartbeat classification using feature selection driven by database generalization criteria. *IEEE Trans Biomed Eng*, 58(3):616–25, 2011.

[96] Dua S. et al. Novel classification of coronary artery disease using heart rate variability analysis. *Journal of Mechanics in Medicine and Biology*, 12:1240017, 2012.

[97] Rajeswari K. et al. Feature selection in ischemic heart disease identification using feed forward neural networks. *Procedia Engineering*, 41:1818–1823, 2012.

[98] Kukar M.et al. Analysing and improving the diagnosis of ischaemic heart disease with machine learning. *Artif Intell Med*, 16(1):25–50, 1999.

[99] Jekova I. et al. Pattern recognition and optimal parameter selection in premature ventricular contraction classification. In *Computers in Cardiology*, 2004.

[100] Behadada O. et al. An interpretable classifier for detection of cardiac arrhythmias by using the fuzzy decision tree. *Artif Intell Res*, 2:45–58, 2013.

[101] Wieben O. et al. Classification of premature ventricular complexes using filter bank features, induction of decision trees and a fuzzy rule-based system. *Med Biol Eng Comput*, 37(5):560–5, 1999.

[102] Kannathal N. et al. Cardiac state diagnosis using adaptive neuro-fuzzy technique. In *IEEE Eng Med Biol Soc*, 2005.

[103] Lin Y. Patterns recognition using multilayer perceptron and classification tree. In *Proceedings of International Medical Informatics Symposium*, 2007.

[104] Park J. et al. Pchd: personalized classification of heartbeat types using a decision tree. *Comput Biol Med*, 54(79-88), 2014.

[105] Khazaee A. et al. Classification of electrocardiogram signals with support vector machines and genetic algorithms using power spectral features. *Biomedical Signal Processing and Control*, 5(252-263), 2010.

[106] Acır N. A support vector machine classifier algorithm based on a perturbation method and its application to ECG beat recognition systems. *Expert Systems with Applications*, 31(1):150–158, 2006.

[107] Jiang W. at al. Block-based neural networks for personalized ecg signal classification. *IEEE Trans Neural Netw*, 18(6):1750–61, 2007.

[108] Acharya R. et al. Classification of cardiac abnormalities using heart rate signals. *Med Biol Eng Comput*, 42(3):288–93, 2004.

[109] Thirugnanam M. et al. Hybrid feature extraction and stacking based ensemble classifier model for cardiomiopathy classification. *International Journal of Advanced Science and Technology*, 29(3):1396–1413, 2020.

[110] Mahesh V. et al. Cardiac disease classification using heart rate signals. *Int J Electron Healthc*, 5(3):211–30, 2009.

[111] Breiman L. et al. *Classification-and-Regression-Trees-Wadsworth-Statistics-Probability*. Chapman and Hall/CRC, 1983.

[112] Roche F. et al. Screening of obstructive sleep apnea syndrome by heart rate variability analysis. *Circulation*, 28(100):1411–5, 1999.

[113] Mateo J. Improved heart rate variability signal analysis from the beat occurrence times according to the ipfm model. *IEEE Transactions on Biomedical Engineering*, 47(8):985–996, 2000.

[114] Berger R.D. et al. An efficient algorithm for spectral analysis of heart rate variability. *IEEE Trans Biomed Eng*, 33(3):900–4, 1986.

[115] Moody G.B. Spectral analysis of heart rate without resampling. *Proceedings of Computers in Cardiology Conference*, 1993.

[116] Abeysekera S.S. et al. Spectral information changes in obtaining heart rate variability from tachometer R-R interval signals. *Crit Rev Biomed Eng*, 28(1-2):149–155, 2000.

[117] Fornasa E. et al. Hrv spectral and fractal analysis in heart failure patients with different aetiologies. In *Computing in Cardiology*, 2014.

[118] Welch P. The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, 15:70–73, 1967.

[119] Rangayyan R.M. *Biomedical Signal Analysis*. IEEE Press Series on Biomedical Engineering, 2015.

[120] Pagani M. et al. Power spectral density of heart rate variability as an index of sympatho-vagal interaction in normal and hypertensive subjects. *J Hypertens Suppl*, 2(3):S383–5, 1984.

[121] Takens F. *Detecting Strange Attractors in Turbulence. Lecture Notes in Mathematics*. Springer, 2006.

[122] Kamen P.W. et al. Poincaré plot of heart rate variability allows quantitative display of parasympathetic nervous activity in humans. *Clin Sci*, 91(2):201–8, 1996.

[123] D'Addio G. et al. Reproducibility of short- and long-term Poincare plot parameters compared with frequency-domain HRV indexes in congestive heart failure. In *Computing in Cardiology*, 1998.

[124] Huikuri H.V. et al. Abnormalities in beat-to-beat dynamics of heart rate before the spontaneous onset of life-threatening ventricular tachyarrhythmias in patients with prior myocardial infarction. *Circulation*, 15(93):1836–44, 1996.

[125] Tulppo M.P. et al. Quantitative beat-to-beat analysis of heart rate dynamics during exercise. *Am J Physiol*, 271:H244–52, 1996.

[126] Mandelbrot B.B. et al. Fractional brownian motions, fractional noises and applications. *SIAM Rev*, 10(4):422–437, 1968.

[127] Accardo A. et al. Use of the fractal dimension for the analysis of electroencephalographic time series. *Biol Cybern*, 77(5):339–50, 1997.

[128] Raghavendra B.S. et al. A note on fractal dimensions of biomedical waveforms. *Comput Biol Med*, 39(11):1006–12, 2009.

[129] Paramanathan P. et al. Application of fractal theory in analysis of human electroencephalographic signals. *Comput Biol Med*, 38(3):372–8, 2008.

[130] Esteller R. et al. A comparison of waveform fractal dimension algorithms. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 48(2):177–183, 2001.

[131] Higuchi T. Approach to an irregular time series on the basis of the fractal theory. *Physica D: Nonlinear Phenomena*, 31(2):277–283, 1988.

[132] Cusenza M. et al. Relationship between fractal dimension and power-law exponent of heart rate variability in normal and heart failure subjects. In *Computing in Cardiology*, 2010.

[133] Rankine L. et al. A nonstationary model of newborn eeg. *IEEE Trans Biomed Eng*, 54(1):19–28, 2007.

[134] Haykin S. *Neural Networks: A Comprehensive Foundation*. Springer, 1994.

[135] Bishop C.M. Neural networks and their applications. *Review of Scientific Instruments*, 65(6), 1994.

[136] Goh A.T. Back-propagation neural networks for modeling complex systems. *Artif Intell Eng*, 9(9):143–151, 1995.

[137] C. Bishop. *Neural networks for pattern recognition*. Clarendon Press, 1996.

[138] Russell S.J et al. *Artificial Intelligence: A Modern Approach-Second Edition*. Prentice Hall, 2003.

[139] Galvao R.K.H. et al. *3.05 - Variable Selection*. Comprehensive Chemometrics Chemical and Biochemical Data Analysis, 2015.

[140] I.T. Jolliffe. *Principal Components Analysis*. Springer series in statistic, 2012.

[141] Dulli S. et al. *Data mining*. Springer, 2009.

[142] K. Hajian-Tilaki. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian J Intern Med*, 4(2):627–35, 2013.

[143] World Health Organization. *Prevention of Cardiovascular Disease Guidelines for assessment and management of cardiovascular risk. Guidelines for assessment and management of cardiovascular risk*. WHO Library Cataloguing, 2007.