Mestrado Integrado em Engenharia Biomédica

Faculdade de Ciência e Tecnologia da Universidade de Coimbra

Validation of non – invasive electromechanical sensors for cardiac monitoring

Clinical trials and implementation of data mining techniques

João Manuel de Olim Perestrelo Borba

Coimbra, September 2012

Validation of non – invasive electromechanical sensors for cardiac monitoring

Clinical trials and implementation of data mining techniques

Scientific Advisors:

PhD Professor Carlos M. B. A. Correia

PhD João Manuel Rendeiro Cardoso

Scientific Supervisor:

MsC Vânia Maria Gomes de Almeida

Dissertation presented to the University of Coimbra to complete the necessary requirements to obtain the MSc degree in Biomedical Engineering

João Manuel de Olim Perestrelo Borba

Coimbra, September 2012

This work is funded by FEDER, through the *Programa Operacional Factores de Competitividade* - COMPETE and by National funds through FCT – *Fundação para a Ciência e Tecnologia in the frame of “Centro de Instrumentação – Unidade217”*.

“The real voyage of discovery consists not in seeking new landscapes, but in having new eyes.”

- *Marcel Proust*

Ao meu pai, que continua
a olhar por mim.

Acknowledgements

From all the lessons that life has taught me so far, there is one that will surely endure through my personal and professional future: that with hard work and perseverance, there is no goal that cannot be achieved. However, I would have not succeeded without the help and support of everyone who accompanied me along the road.

First of all, I would like to thank my family, especially my mother, sister and grandparents, for all their constant love and support throughout my studies. A special thanks to my dear Carolina, for all her love and for helping me achieve a piece of mind when I most need it.

I would like to express my great appreciation to Prof. Dr. Carlos Correia and Dr. João Cardoso, for all their guidance and support. I am grateful to my supervisor Vânia Almeida, for all her encouragement and mentoring, which was very important in achieving success in this project. I thank my work colleagues Inês Santos, André Cortez, Mariana Sequeira, Anurati Saha, Pedro Santos and Pedro Vaz, for all their help and companionship during these past ten months.

I also want to thank to my apartment colleagues, Tiago and André, for their close friendship. Last, but not least, a special thanks to my university colleagues, Sérgio Pinto and Susete Neiva, for helping me in overcoming all the obstacles along the course.

To Nuno, a friend whom I will never forget.

Thank you all.

- *João Borba*

Abstract

Nowadays, arterial stiffness assumes special importance as the result of being a marker of cardiovascular diseases (CVD), which are the leading cause of disability and worldwide. The development of diagnostic tools capable of performing an early and precise quantification of pathologic states such as arterial stiffness presents itself as a global strategy to reduce cardiovascular (CV) morbidity and mortality.

A non – invasive (PZ) technology for arterial distension waveform (ADW) monitoring in the carotid artery has been successfully developed and tested in the past few years. This piezoelectric (PZ) device allows the extraction of clinically important information concerning arterial stiffness, presenting itself as a practical solution in premature CV risk assessment.

This project consisted not only in the first clinical trials of the previously developed device with the performance of repeatability tests, but also in the application of innovative data mining tools such as classification and clustering approaches, with the objective of developing innovate decision support systems for CV risk estimation. Finally, a case study was also carried out in patients that suffered from severe stenosis to prove the usefulness of this technology.

Excellent repeatability results between trials were obtained. Furthermore, the ability to detect physiological variations after surgical procedures has demonstrated the clinical feasibility of this equipment. Data mining methodologies have also shown their effectiveness in premature CV risk determination.

Keywords: *Arterial Stiffness, Arterial Distension Waveform, Piezoelectric Sensor, Clinical Trials Data Mining.*

Resumo

Atualmente, a rigidez arterial assume especial importância pelo facto de ser um marcador de doenças cardiovasculares, que são a principal causa de incapacidade e morte no mundo. O desenvolvimento de ferramentas de diagnóstico que são capazes de realizar uma quantificação exata e prematura de estados patológicos como a rigidez arterial apresenta-se como uma estratégia global para reduzir a morbidade e mortalidade cardiovascular.

Com o intuito de monitorizar a onda de distensão arterial na carótida, uma tecnologia não – invasiva foi desenvolvida e testada com sucesso durante os últimos anos. Esteve dispositivo piezoelétrico permite a extração de informações clinicamente importantes sobre a rigidez arterial, apresentando-se como uma solução prática na avaliação do risco cardiovascular prematuro.

Este projeto consistiu não só no início dos primeiros testes clínicos do dispositivo previamente desenvolvido com a realização de testes de repetibilidade, mas também na aplicação de ferramentas inovadoras de mineração de dados através de abordagens classificativas e de agrupamento. Por último, um caso de estudo foi realizado em doentes com estenose severa de forma a provar a utilidade desta tecnologia.

Foram obtidos excelentes resultados em termos da repetibilidade entre ensaios consecutivos. Além disso, a capacidade de detetar variações fisiológicas após procedimentos cirúrgicos demonstrou a aplicabilidade clínica deste equipamento. As metodologias de mineração de dados também mostraram a sua eficácia na determinação prematura de risco cardiovascular.

Palavras – Chave: *Rigidez Arterial, Onda de Distensão Arterial, Sensor Piezoelétrico, Ensaios Clínicos, Mineração de Dados.*

Contents

List of Figures

List of Tables

Acronyms

ADW	Arterial Distension Waveform
AGE	Advanced Glycation Endproducts
AIx	Augmentation Index
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
BMI	Body Mass Index
BPM	Beats Per Minute
CR	Coefficient of Repeatability
CRP	C – Reactive Protein
CV	Cardiovascular
CVD	Cardiovascular Diseases
CVN	Cross – Validation
DBP	Diastolic Blood Pressure
DW	Dicrotic Wave
DW_A	Dicrotic Wave Amplitude
DW_T	Dicrotic Wave Time
EM	Expectation - Maximization
FN	False Negative
FP	False Positive
FWHM	Full Width at Half Maximum
GUI	Graphical User Interface
HR	Heart Rate
IQR	Interquartile Range
IREP	Incremental Reduced Error Pruning
KDD	Knowledge Discovery from Data
KS	Kolmogorov - Smirnov
LNN	Linear Neural Network
MLP	Multi – Layer Perceptron
MRI	Magnetic Resonance Imaging
NIST	National Institute of Standards and Technology
PI	Point of Inflection
PP	Pulse Pressure
PVC	Polyvinyl Chloride

PWV	Pulse Wave Velocity
PZ	Piezoelectric
RBF	Radial Basis Function
RIPPER	Repeated Incremental Pruning to Produce Error Reduction
RMSE	Root Mean Square Error
RMSSD	Root Mean Square of Successive Differences
RP_A	Reflection Point Amplitude
RP_T	Reflection Point Time
SBP	Systolic Blood Pressure
SD	Standard Deviation
SLP	Single – Layer Perceptron
SNR	Signal-to-Noise Ratio
SP	Systolic Peak
SP_A	Systolic Point Amplitude
SP_T	Systolic Point Time
FWHM	Full Width at Half Maximum
TN	True Negative
TP	True Positive
WHO	World Health Organization

List of Original Papers

Parts of this dissertation have been published in the following papers:

I. Validation of a waveform delineator device for cardiac studies: repeatability and data mining analysis, V. G. Almeida, **J. Borba**, T. Pereira, H.C. Pereira, J. M. R. Cardoso, C. Correia, 2012. Accepted to 2nd National Meeting of Bioengineering, 23 – 25 February, Coimbra, Portugal 4 pp.

II. Indices and Repeatability Tests of Cardiovascular Function Performed on the Arterial Distension Waveform – Case Study: Angiography Intervention, V. G. Almeida, **J. Borba**, H.C. Pereira, T. Pereira, J. Cardoso, C. Correia, 2012. Submitted to BIOSTEC 2013 – Biosignals (6th International Joint Conference on Biomedical Engineering Systems and Technologies), 11 – 14 February, Barcelona, Spain, 8 pp.

III. Data Mining Based Methodologies for Cardiac Risk Patterns Identification, V. G. Almeida, **J. Borba**, T. Pereira, H. C. Pereira, J. Cardoso, C. Correia, 2012. Submitted to BIOSTEC 2013 – Bioinformatics (6th International Joint Conference on Biomedical Engineering System and Technologies), 11 – 14 February, Barcelona, Spain, 8 pp.

In Paper I, results for elementary repeatability measurements and data mining based on classifications procedures are presented and discussed. ***Paper I is presented in Appendix E.***

Paper II focuses not only in the repeatability analysis between sessions and between different carotid sites, but also in a stenosis case study with the objective of demonstrating the feasibility of the sensor in detecting distinct clinical patterns.

Paper III demonstrates the potentialities of data mining based methodologies in assessing cardiovascular risk.

1. Introduction

"The best way to predict the future is to invent it"

- Alan Kay

1.1 Motivation

According to World Health Organization (WHO), cardiovascular diseases (CVD) represent 31% of all global deaths, making them the leading cause of death and disability worldwide (Figure 1.1). In 2008, 17.3 million people died because of CVD. Moreover, it has been revealed that CVD mortality had increased at an impressive fast rate in low- and middle-income countries during the past twenty years. The global strategies to reduce incidence, morbidity and mortality of CVD comprise an efficient identification and reduction of CVD risk factors and their determinants and the elaboration of moderate-cost diagnostic tools that can be used in early stages of CVD development [1].

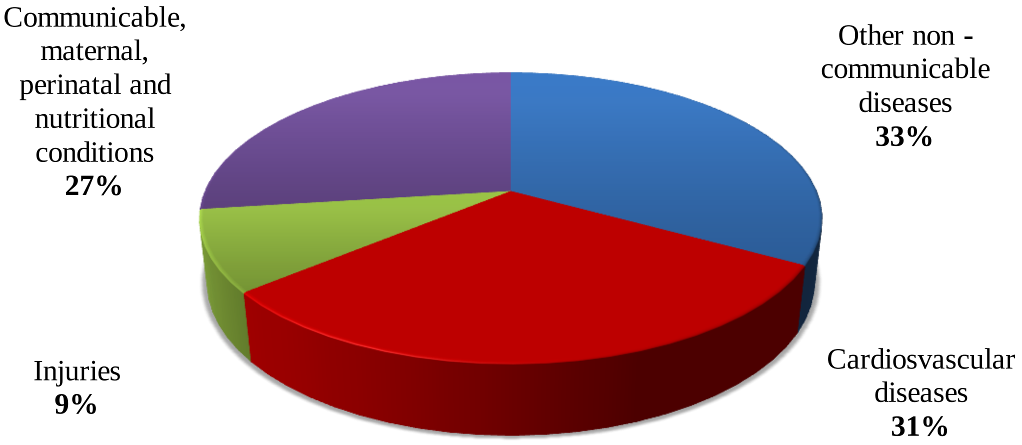


Figure 1.1 – Distribution of major causes of death in the world according to a 2008 report from WHO [1];

Current investigations are focused on non-invasive measures of arterial function, which are associated with and are prognostic of CVD in the general adult and elderly population [2-5]. The presumption that arterial stiffness is a marker of cardiovascular (CV) events has gained support over the last years due to repeated demonstration of its predictive value for myocardial infarction, stroke and CV death [6-9]. Therefore, it is believed that early quantification of arterial stiffness can ultimately reduce CV morbidity and mortality.

Much interest has been paid to the arterial distension waveform (ADW), which contains a vast amount of pathophysiological information hidden in its morphology, and since it is possible to non – invasively extract ADW parameters that reveal important clinical information concerning arterial stiffness, the development of efficient non – invasive ADW acquisition modules gained significant relevance. Among diverse non – invasive acquisition techniques, the use of piezoelectric (PZ) sensors in ADW measurements has been reported by several authors as having a good performance, as result of their high sensitivity and high signal-to-noise ratio (SNR) [10-12]. Previous established developments have contributed for a non – invasive acquisition device that can be a convenient and suitable solution to assess the hemodynamic condition [13, 14].

The analysis and exploration of large databases is an important issue in the development of subsequent algorithms and devices. Therefore, it is critical to expand a database which contains subjects not only with cardiac pathologies, but healthy subjects, as well. To successfully test and validate this non – invasive acquisition device, clinical validation is of crucial importance. Repeatability tests must be performed [15], and so, it is necessary to collect a significant number of subjects with and without CVD to determine the precision of a specific method. Nowadays, the Bland – Altman method is a respected and efficient “gold – standard” for repeatability assessment [16].

On other hand, with the recent emergence of computer – aided diagnosis technologies, new insights for the development of innovative methods for health professional assistance were uncovered. By recognizing relationships and patterns in huge amounts of database, data mining techniques can be decisive in discovering new biomedical and healthcare knowledge for clinical and administrative decision making. In other words, it is expected from data mining techniques their becoming as the keystone of future healthcare decision support systems [17-20].

1.2 Previous work

The present work is a continuity project with the ultimate objective of developing a highly functional non-invasive device that is capable of efficiently describing and evaluating the CV condition.

The work that fully motivated the use of alternative instrumental methods for non – invasive hemodynamic characterization was Pereira’s MsC dissertation, which was based on the use of accelerometery concepts for pulse wave velocity (PWV) assessment [21]. Almeida’s and Pereira’s MsC dissertations [22, 23] focused in the use and development of PZ sensors for ADW acquisition, and also contributed with the development of algorithms capable of rendering information about hemodynamic parameters. Afterwards, important hardware, firmware and software advancements were accomplished [13, 14], and more recently, Vieira's MsC dissertation constructed a database and improved ADW signal processing routines even further [24, 25]. A preliminary data mining approach was also performed, and proved that data mining classification techniques are an interesting tool in the study of arterial stiffness related patterns. [26].

1.3 Objectives

After considering the previous accomplishments, the proposed objectives for this work are:

- Database expansion of the number of subjects in the database.
- The validation of the previously developed ADW acquisition device using repeatability measures. Clinical trials will be performed to assess the repeatability of the previously developed non – invasive system between trials, between left and right carotid artery and between two successive months. Some of the clinical trials will be performed in collaboration with the *Centro Hospitalar e Universitário de Coimbra.*
- The use of data mining techniques as innovative decision support systems regarding CVD, which includes:
    - Classification techniques, for the development of diagnostic procedures to determine the CV condition of an undiagnosed subject. Focus will be given to artificial neural network (ANN) classifiers, since they were not used in previous works, and already revealed themselves as very effective in the CVD context. [27-31]. A multiple classifier methodology that is mainly based in the method of Gorunescu *et al.* (2011) [32] will be tested as an innovative diagnostic tool.
    - Clustering techniques, to find and distinguish different risk groups in terms of future CVD development in a healthy dataset. Shah et al. (2008) [20] hypothesized that categorizing a youth based dataset into low- and high – risk groups can be a simple and effective tool for evaluating the risk of developing CVD in young adults.

1.4 Project team composition

This work was developed at *Grupo de Electrónica e Instrumentação* (GEI), one of the research groups of *Centro de Instrumentação* (CI) of the University of Coimbra (UC). GEI is especially dedicated to the study and development of instrumentation and dedicated software for biomedical and physics applications, in close partnership with *Intelligent Sensing Anywhere* (ISA).

Table 1.1 shows an overview of the main contributors of this project and their associated staff and students.

Table 1.1 – Project team members;

Team Members	*Contribution or main area of research*
PhD Professor Carlos M. B. A. Correia **PhD João Manuel Rendeiro Cardoso**	*Scientific and technical advisors*
PhD Student Vânia Maria Gomes de Almeida	*Scientific and technical supervisor / Study of hemodynamic parameters*
MsC Student João Manuel de Olim Perestrelo Borba	*Student responsible for the development of the project*

1.5 Contents by chapter

This dissertation is divided in eight chapters, excluding the references. In the first chapter **(Introduction)**, the motivation of this thesis, its main contributions and the project team composition are introduced. Lastly, the chapter-by-chapter structure of the thesis is briefly explained.

In the second chapter **(Theoretical Background)**, the main physiological concepts behind this dissertation are presented, with a special focus in defining arterial stiffness and its determinant factors in the development of CVD. Distinctive methods for non – invasive arterial stiffness assessment are also described, and the angioplasty procedure is detailed as well.

The third chapter (**Clinical Trials – Repeatability)** approaches on how to evaluate the inherent precision of an equipment using general descriptive statistics, correlation techniques, and the Bland – Altman method. Some important studies related with the repeatability of non – invasive CV assessment are described in the last sub – chapter.

The fourth chapter (**Data Mining)** focuses in introducing data mining and its importance in the new century. Data pre – processing, data classification and data clustering compose the other sub – chapters, and are rigorously described.

In the fifth chapter (**Hardware & Software)**, the state of art of all the hardware and software that were previously developed is summarily described. Graphical user interface (GUI) improvements are also referred.

The sixth chapter **(Methodology)** centers on the work – oriented methodology, in other words, all the procedures to obtain results are fully described, including data acquisition and data pre – processing setups, feature selection and database characterization. The creation and development of a Bland – Altman GUI is another point of interest. Four different datasets, as each dataset will be used for a specific study. Dataset I is used for assessing the repeatability of the non – invasive system, dataset II is used for a case study with the objective of proving the efficiency of the PZ probe in detecting physiological changes between two different conditions. Dataset III and dataset IV focuses on data mining methodologies (classification and clustering, respectively).

In the seventh chapter **(Results & Discussion)**, all the obtained results are presented and discussed for each of the datasets. In the final sub – section of each dataset, results are extensively discussed.

In the eighth and final chapter (**Conclusion & Future Work)**, conclusions are assessed from the developed work, and the main contributions are correlated with the initially proposed objectives. Various suggestions and guidelines for a possible future work are presented.

Table 1.2 – Gantt diagram of the project tasks;

Id	***Task Name***	***Start***	***Finish***	***2011***				***2012***							
				Set	*Out*	*Nov*	*Dez*	*Jan*	*Fev*	*Mar*	*Abr*	*Mai*	*Jun*	*Jul*	*Ago*
1	State of the art review	12-09-2011	16-12-2011	■	■	■	■								
2	Elementary repeatability measurements	26-09-2011	16-12-2011	■	■	■	■								
3	Dataset II arrangement	21-11-2011	02-12-2011			■									
4	Data processing and statistical analysis for elementary repeatability measurements	02-12-2011	16-12-2011				■								
5	Writing of paper I – "Validation of a waveform delineator for cardiac studies: repeatability and data mining analysis"	02-01-2012	16-01-2012					■							
6	Data mining: classification procedures (dataset II)	19-01-2012	17-02-2012					■	■						
7	Poster for 2nd National Meeting of Biomedical Engineering	20-02-2012	23-02-2012						■						
8	Final repeatability measurements (dataset I)	27-02-2012	23-03-2012						■	■					
9	First massive data acquisition (dataset III)	16-03-2012	30-03-2012							■					
10	Development of the Bland – Altman GUI	09-04-2012	17-04-2012								■				
11	Data processing and statistical analysis for final repeatability measurements (dataset I)	18-04-2012	04-05-2012								■	■			
12	Second massive data acquisition (dataset III)	30-04-2012	04-05-2012								■	■			
13	Data mining: clustering procedures (dataset III)	14-05-2012	01-06-2012									■			
14	Preliminary project presentation	04-06-2012	06-06-2012										■		
15	Writing of paper II – "Indices and Repeatability Tests of Cardiovascular Function Performed on the Arterial Distension Waveform – Case study: Angiography Intervention"	11-06-2012	29-06-2012										■		
16	Writing of paper III – "Data Mining Based Methodologies for Cardiac Risk Patterns Identification"	11-06-2012	29-06-2012										■		
17	Thesis writing	02-07-2012	31-08-20122											■	■

2. Theoretical Background

In this chapter, the main physiological concepts of this thesis are discussed. An overview of the CV system and the role of arterial stiffness in the development of CVD are given, and several methods and indexes for non – invasive assessment of arterial stiffness are explained. The angioplasty procedure is also detailed.

2.1 Cardiovascular system

The human physiology possesses an effective CV system, which uses blood as a working fluid, and consists of the heart, arteries, veins, capillaries and lymphatic vessels. The CV system has three important functions [33]:

- Supply oxygen and nutrients to body tissues
- Remove carbon dioxide and other wastes from the body
- Regulate temperature

2.1.1 The heart

The heart is a vital organ of our physiological system, anatomically located between the third and sixth ribs in the central portion of the thoracic cavity. But actually, the heart is composed by two detached pumps, separated by a tough muscular wall (interventricular septum): a *right heart*, which pumps blood through the lungs, and a *left heart* that pumps blood through the peripheral organs. Each heart is a pulsatile two – chamber pump, composed of an atrium and a ventricle. Each atrium is an essential pump that moves blood into the ventricle. The ventricles provide the major source of power that propels the blood through the *pulmonary circulation* (right ventricle) or through the *systemic circulation* (left ventricle) [34]. Check valves between each set of upper and lower chambers ensures that the blood moves in only one direction and enables the pressure in the aorta to be much higher than the pressure in the lungs. This restricts the blood from flowing backwards [34]. A structure of the heart and the course of the blood flow through the heart's chambers can be found on figure 2.1.

The pressure values differ between the left pump and the right pump. Because of the anatomic proximity of the heart to the lungs, the right side of the heart does not have to work very hard to drive blood through the pulmonary circulation, so it functions at a low – pressure (< 40 mmHg gauge). In contrast, the left side of the heart does most of its work at a high pressure value (up to 140 mHg gauge or more) to drive blood through the whole systemic circulation [33, 34].

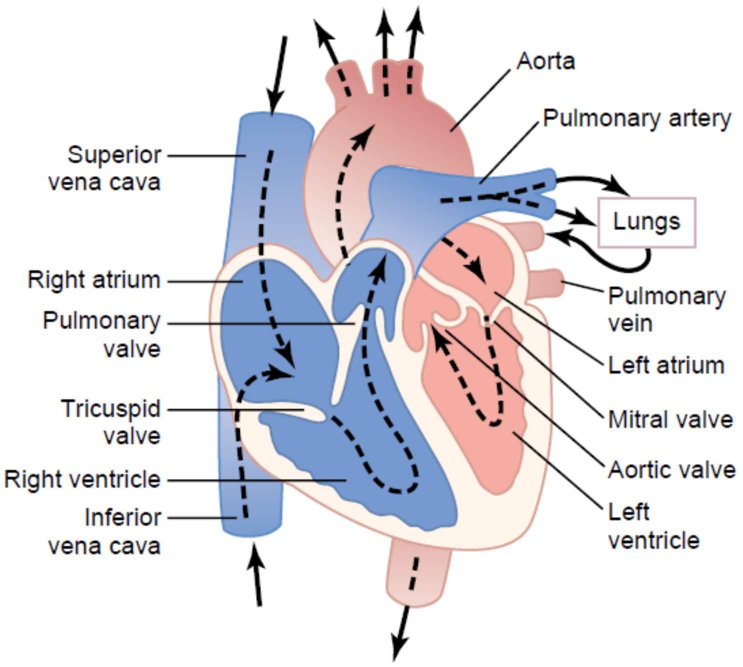


Figure 2.1 – Structure of the heart and the course of the blood flow. From [34];

2.1.2. Common carotid artery

The left and right common carotid arteries provide the major source of blood to the head, neck and brain. Both the right and left common carotid arteries branch into internal and external carotid arteries. The internal carotid artery supplies oxygenated blood to the brain and eyes, while the external carotid artery provides oxygenated blood to the throat, neck glands, face, scalp, mouth and tongue. Both left and right common carotid arteries differ with respect to their origins. In the left common carotid, the artery comes directly from the arch of the aorta in the superior mediastinum. The right common carotid artery arises from the brachiocephalic artery as it passes behind the sternoclavicular joint [34, 35].

2.2 Arterial stiffness

Arterial stiffness measures the rigidity in the arterial wall or, in other words, the arteries capacity of expanding and contracting during the cardiac cycle [4]. Nowadays, there is a scientific consensus regarding the importance of arterial stiffness in the development of CVD, being the most important parameter of increasing systolic and pulse pressure (PP) [4, 6].

2.2.1 Arterial structure

All the arteries are composed of three main layers: intima, media and externa (Figure 2.2). The intima is the innermost layer, a single layer of endothelial cells and associated connective tissue. The middle layer is the media, and is composed of a specific amount of elastic and smooth muscle fibers, which varies depending on the size and location of the artery in the arterial tree. The outer layer is the externa (or adventitia), which serves as a connective tissue is largely composed of collagen fibers [36, 37].

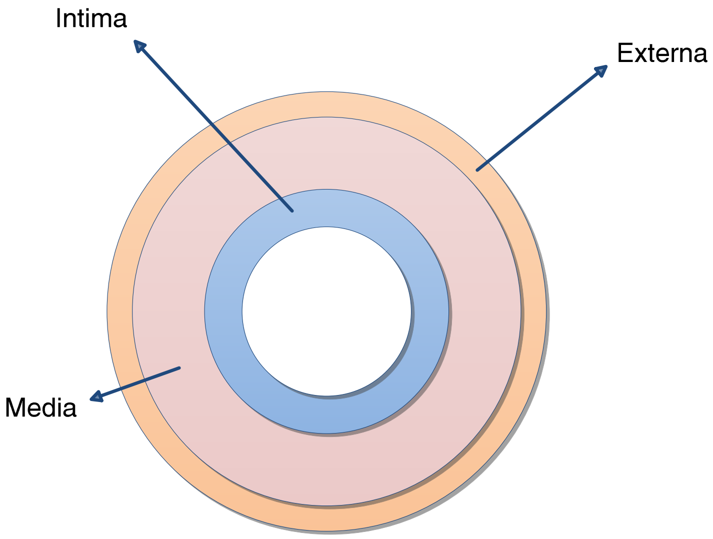


Figure 2.2 – Representation of the arterial structure components: intima, media and externa. Adapted from [5];

2.2.2 Arterial stiffness mechanisms

The most accepted model of the arterial tree is the propagative model, which consists of a viscoelastic tube whose distributed elastic properties permit generation of a forward pressure wave which travels along the tube and whose numerous branch points and high level of resistance of tube's end generate reflected waves [38]. If we would just consider a viscoelastic tube without reflection sites, the pressure wave would be progressively attenuated, with an exponential decay along the tube. In contrast, a pressure wave that propagates along a viscoelastic tube with numerous branches is progressively amplified due to wave reflections, from central to distal conduit arteries. Because of this, the amplitude of the pressure wave is higher in peripheral arteries than in central arteries. Therefore, it is not accurate to use brachial (upper arm artery) PP as a perfect substitute for aortic or carotid PP, especially in younger subjects [6].

2.2.3 Proximal and distal arterial stiffness

The elastic properties of conduit arteries vary along the arterial tree. Proximal arteries are more elastic and, in contrast, distal arteries are stiffer. The elasticity of the proximal large arteries is the result of the high elastin to collagen ratio in their walls, which progressively declines toward the periphery (Figure 2.3).

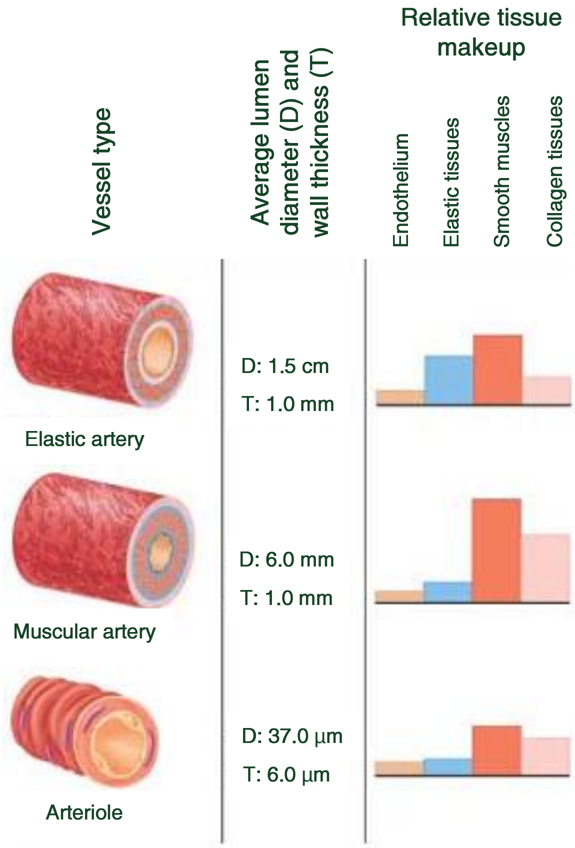


Figure 2.3 – Comparison of different arterial vessel types, regarding their average lumen diameter, wall thickness, and relative tissue makeup. Adapted from [37];

2.2.4 Associated pathophysiological conditions

Various reversible and irreversible pathophysiological conditions are associated with an increase in arterial stiffness. Those conditions are expressed on Table 2.1.

Table 2.1 – Pathophysiological conditions that affect arterial stiffness. Adapted from [6];

Aging	**CV risk factors**	**CV diseases**
Other physiological conditions	Obesity	Coronary heart disease
Low birth weight	Smoking	Congestive heart failure
Menopausal status	Hypertension	Fatal stroke
Lack of physical activity	Hypercholesterolaemia	**Primarily non – CV diseases**
Genetic background	Impaired glucose tolerance	End – stage renal disease
Parental history of hypertension	Metabolic syndrome	Moderate chronic kidney disease
Parental history of diabetes	Diabetes type 1 and 2	Rheumatoid arthritis
Parental history of myocardial infarction	Hyperhomocyteinaemia	Systemic vasculitis
Genetic polymorphisms	High C – Reactive Protein (CRP)	Systemic lupus erythematosus

2.2.4.1 Aging

Age is the most important determinant of arterial stiffness, as stiffening of large arteries is a consequence of the normal aging process. The aging process in the arterial tree is not homogeneous, as the elastic properties of distal (and more muscular) arteries change little with age.

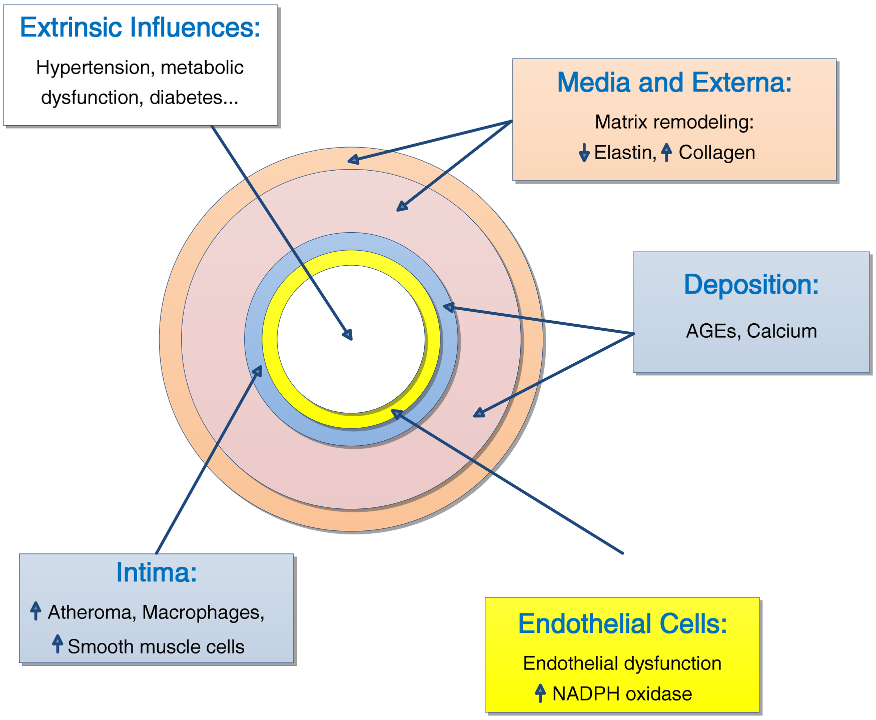


Figure 2.4 – Causes of arterial aging in a common elastic artery. Adapted from [5];

The main structural change with aging is medial degeneration, which leads to progressive stiffening of large elastic arteries. Longstanding arterial pulsation in the central artery has a direct effect on the structural matrix proteins, collagen and elastin in the arterial wall, disrupting muscular attachments and causing elastin filaments to fatigue and fracture. Other main causes include changes in the vascular smooth muscle cells also mediate aging – associated vascular stiffness, endothelial dysfunction triggered by a decrease in anti - oxidative capacity and an increase in oxidative stress, accumulation of advanced glycation endproducts (AGE) on the proteins and calcium deposition in the arterial wall. Extrinsic factors that can eventually appear with advanced aging may also play a role [5].

2.3 Arterial distension waveform

Clinically relevant information can be found in the ADW morphology. Numerous methods for ADW acquisition and subsequent analysis are used nowadays, such as invasive catheterization and, more recently, non – invasive applanation tonometry.

The ADW can be acquired at a central or peripheral level. The measurement at peripheral zones (such as radial, brachial or femoral artery) uses a transfer function to reconstruct aortic waveform which decreases the accuracy of data. In contrast, the measurement at the central artery surrogates the true load imposed to the left ventricle and central large artery walls. While it requires a higher degree of technical expertise, a transfer function is not necessary, thus increasing data precision. The central artery measurement is usually done in the carotid arteries, as they are very close to central artery, so their waveforms are equivalent [22].

2.3.1 ADW morphology

Considering an acquisition at a central level, the ADW is composed by two main components: a forward incident wave, caused by left ventricular contraction and ejection of blood into the arterial tree; a backward reflected wave from the periphery that returns to the heart due to arterial tree branch points or sites of impedance mismatch [39].

2.3.1.1 Incident wave

The incident wave occurs due to the capacitive characteristics of the ascending aorta segment, after left ventricular blood ejection. Its characteristics depend largely on the left ventricular ejection and aorta stiffness [6, 39].

Usually, the incident wave has two points of interest that can be observed on the ADW. The first point is the systolic peak (SP), and corresponds to the highest pressure value of the ADW. The other zone of interest is the incisura (also known as dicrotic wave – DW), which corresponds to an increase of the aortic pressure along the ascending aorta after the closure of the aortic valve. The incisura phenomenon can be used to obtain systolic duration [40].

2.3.1.2 Reflected wave

The characteristics of the backward reflected wave are influenced by reflections coefficients values, sites of reflection points and elastic properties of the arterial tree [39].

On an ADW, a point of inflection (PI) that directly corresponds to the backward reflected wave can usually be observed (except in one case, more on that on the following sub - chapter 2.3.2).

2.3.2 ADW types

According to Murgo et al. (1986) [41], four types of ADW can be described (Figure 2.5), and the determinant criterion for wave classification is the location of the reflected wave.

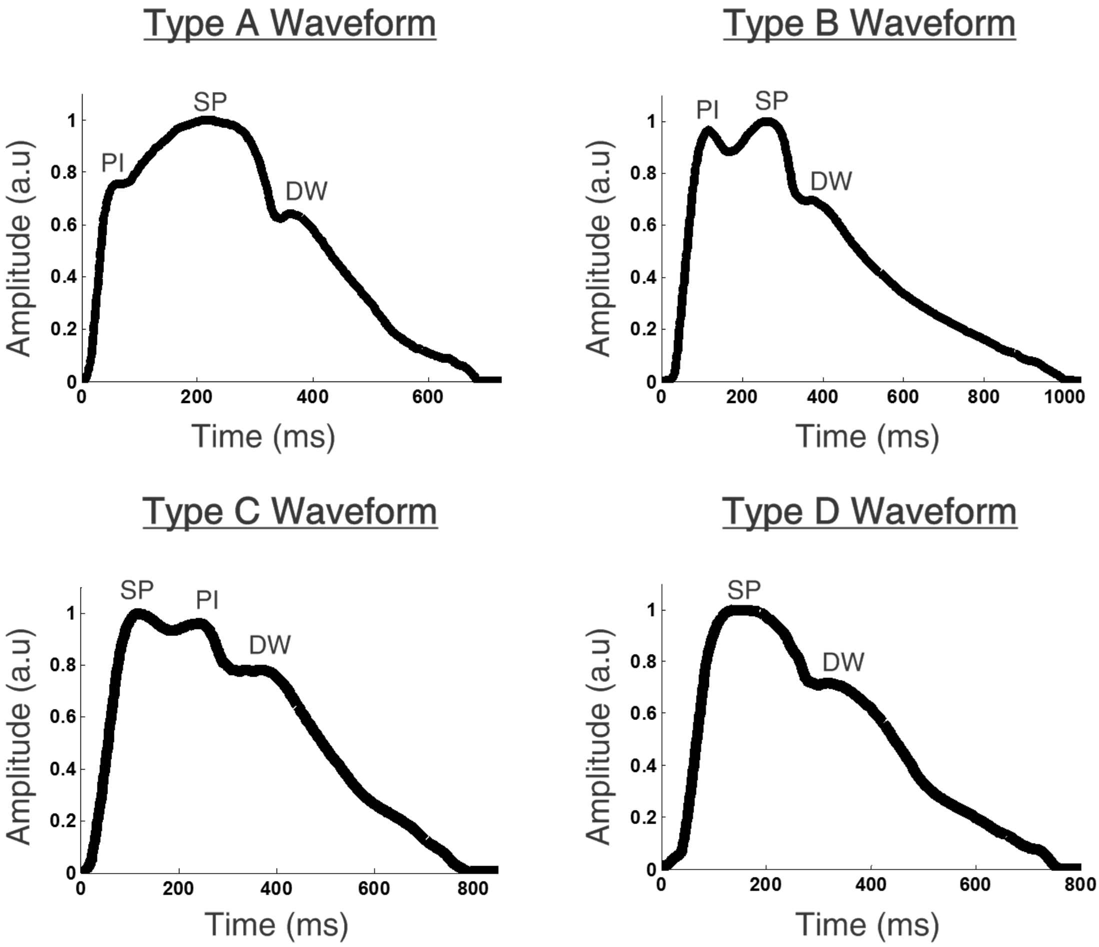


Figure 2.5 – APW classification according to Murgo *et al* (1989) [41]. SP represents the systolic peak, PI is the point of inflection and DW is the dicrotic wave;

2.4 Non – invasive assessment of arterial stiffness

Several hemodynamic parameters can be used to assess arterial stiffness. All require information about simultaneous change in arterial size and arterial pressure in order to quantify the change in arterial stiffness.

2.4.1 Pulse pressure (PP)

PP is the difference between systolic blood pressure (SBP) and diastolic blood pressure (DBP). It is considered as a valuable surrogate marker of arterial stiffness as it depends on cardiac output, large artery stiffness, and wave reflections [36]. A high PP is often a marker that the heart is working harder than the usual to maintain a homeostatic circulation. PP is also known to increase with age [36]. Usually, most PP measures are made from the brachial artery using an oscillometric sphygmomanometer.

PP measurement is very used in clinical setting due to being a very simple and relatively efficient technique. However, PP measurements do not consider changes in volume, and therefore are not true measures of arterial stiffness. Another problem is the amplification of the pressure wave in the periphery [22].

2.4.2 Arterial compliance and distensibility

Arterial compliance (C) is defined as the change in volume for a given change in pressure and arterial distensibility (D) is the compliance divided by the initial volume:

$$C = \frac{\Delta A}{PP} \tag{2.1}$$

$$D = \frac{\Delta A}{A_d PP} \tag{2.2}$$

With *PP* being the pulse pressure and Δ*A* being the pulse cross – sectional area ($\Delta A = A_s - A_d$), with A being the systolic cross-section area and A_d being the diastolic cross-section area.

Arterial compliance and arterial distensibility can be registered by magnetic resonance imaging (MRI), which records the maximum and minimum arterial diameter. The ultrasound technique has the advantage of being non-invasive, but the equipment is costly and hard to expertise [22].

2.4.3 Pulse wave velocity (PWV)

PWV is the speed at which the ADW generated by cardiac contraction travels from the aorta through the arterial tree. Studies have showed that PWV is an independent predictor of CVD while associated with diverse pathophysiological conditions [42].

PWV is measured using applanation tonometry such as the Complior (Colson, Paris, France) and the SphygmoCor (ArtCor, Sydney, Australia). Even though several different measurement sites can be found in the literature, carotid – femoral is the most common pathway to evaluate regional arterial stiffness. Carotid – femoral PWV is also considered by many as the “gold – standard” non-invasive measurement of arterial stiffness [6, 36]. The transit time is measured by two applanation tonometers placed over the peripheral pulse, and the distance between them is estimated by direct superficial measurement, as expressed in the following equation:

$$PWV = \frac{distance}{\Delta t} \tag{2.3}$$

However, PWV estimation can be quite inaccurate unless the artery between the two pulses is in a straight line. Also, it can be particularly difficult to assess PWV in obese patients [24, 27].

According to Moens and Korteweg (1878) [43], the relationship between arterial stiffness and PWV can be described by the following equation:

$$PWV = \sqrt{\frac{E\ .\ h}{2r\rho}} \tag{2.4}$$

Where *E* is the elastic modulus of the vessel wall, *h* is the wall thickness, *r* is the vessel radius and ρ is the blood density (approximately 1.05). It is assumed that there is insignificant change in vessel area. This equation can be alternatively expressed, according to Bramwell and Will (1922) [44] by the following:

$$PWV = \sqrt{\frac{dP\ .V}{dV\ .\rho}} = \sqrt{\frac{1}{\rho D}} \tag{2.5}$$

Where *P* is the pressure, *V* is the volume, ρ is the blood density, D is the volume distensibility of the arterial segment and *dP . V / dV* represents volume elasticity.

2.4.4 Augmentation index (AIx)

The augmentation index (AIx) measures the strength of the reflected wave relative to the total ADW, and is defined as the difference between the second and first peaks (augmentation pressure) expressed as a ratio or percentage of the PP (Figure 2.6) [6, 22]. In this work, we always considered the AIx expressed as a percentage. AIx has been indicated as a surrogate measure of arterial stiffness and, consequently, a marker of CV risk [45]. AIx is also known to increase with aging [46, 47].

The general equation to determine AIx is the following:

$$AIx = \frac{AP}{PP} = \frac{P_1 - P_2}{PP} \tag{2.6}$$

Where *AP* is the augmentation pressure, *PP* is the pulse pressure, and P_1 is the first pressure peak and P_2 is the second pressure peak.

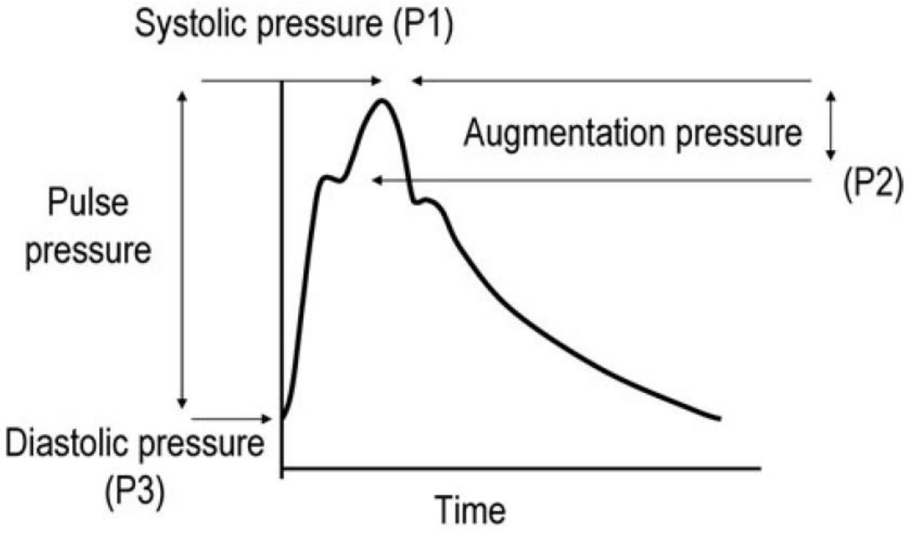


Figure 2.6 – Augmentation pressure as the difference between the systolic pressure and the inflection point pressure. From [6];

The key point to determine AIx is to identify the inflection point. Depending on the location of the reflected wave in the ADW, the first pressure peak can be a reflection point or a systolic point and, consequently, the second pressure peak will be a systolic point or a reflection point, respectively. Table 2.2 explains the differences in AIx calculus and what it could indicate regarding arterial stiffness. As a note of attention, in the following table, P_1 and P_2 were substituted for P_s (systolic pressure) and P_i (pressure at PI or augmentation pressure).

Table 2.2 – Classification of the different APW according to the inflection point position and AIx calculus. P_s is the systolic pressure, P_i is the pressure in the inflection point, and PP is the pulse pressure. Adapted from [27];

ADW Type	*ADW Properties*	*AIx calculus*
A	The inflection point occurs before the systolic peak. The AIx is positive, and indicates high arterial stiffness.	$\frac{P_s - P_i}{PP} \times 100\%$
B	The inflection point occurs shortly before the systolic peak, indicating small arterial stiffness.	$\frac{P_s - P_i}{PP} \times 100\%$
C	The inflection point occurs after the systolic peak. The AIx value is negative, and indicates an elastic and healthy artery.	$\frac{P_i - P_s}{PP} \times 100\%$
D	The inflection point cannot be recognized because the reflected wave arrives in early systole and merges with the incident wave.	*

* As the inflection point cannot be acknowledged, AIx cannot be calculated.

While AIx non – invasive assessment is a very efficient manner of determining local arterial stiffness, it has some minor limitations, including erroneous results when the inflection point is not well identified and the direct influence of blood pressure and heart rate (HR) in AIx, so they must be assessed when measuring AIx [6, 36]

2.4.4.1 Reference values

Few studies regarding AIx reference values can be found in the literature, due to the concept's recent emergence in the scientific community and the inherent difficulty in acquiring a huge sample from the population.

Janner *et. al.* (2010) [46] studied 4561 subjects from The Copenhagen City Heart Study, and calculated reference values of AIx measured by the SphygmoCor. Internally validated AIx reference equations considering age, HR and height were reported for both men and women:

$$AIx = 79.20 + 0.63\ (age) - 0.002(age^2) - 0.28\ (HR) - 0.39\ (height) \tag{2.7}$$

$$AIx = 56.28 + 0.90\ (age) - 0.005(age^2) - 0.24\ (HR) - 0.34\ (height) \tag{2.8}$$

Chung *et al.* (2010) [47] recruited 522 subjects with a mean age of 46.3 years, and measured central and peripheral AIx. They divided their sample in four groups according to decade of age, and concluded the following mean central AIx values in the Korean population:

Table 2.3 – AIx reference values for the Korean population, according to Chung *et al.* (2010) [47];

Age range (years)	*Mean AIx (%)*
≤ 39	23.4
40 – 49	28.9
50 - 59	29.7
≥ 60	34.1

Both studies also assessed that women present higher Aix values than men [46, 47].

2.4.4.2 PWV comparison

Comparisons between AIx and PWV ("gold – standard" method for arterial stiffness measurement) have also been reported. Wimmer *et al.* (2007) [48] reported a medium association between ideal PWV and HR – adjusted AIx (r = 0.371) in chronic kidney disease. Higher associations in women (r = 0.423) relatively to men (r = 0.361) are also documented. Data from a large cohort of healthy individuals in the Anglo – Cardiff Collaborative Trial

(ACCT) showed that central AIx might be a more sensitive marker of arterial aging in young and middle – age individuals (< 50 years) and aortic PWV is more sensitive in the older population (> 50 years) [49].

2.5 Angioplasty with stent placement

Both carotid arteries deliver the required blood for the brain. Sometimes, the blood flow in a carotid artery can become partly blocked due to an artery narrowing (stenosis), increasing the risk of a stroke in short – term. One of the “gold – standard” invasive procedures to treat a narrow or blocked carotid artery is the angioplasty with stent placement.

In this surgical procedure, live angiography is usually the imaging method of choice to visualize the blood vessels. After local anesthesia, a surgical cut is executed next to the groin, and after a filter device is opened above the lesion, the stent is implanted. The balloon is then inflated at the stenosis plaque level, decreasing the stenosis after deflation, and the filter device is kept open as prevention for an eventual embolic trapping. With this surgical technique, the blood flow was restored to the normal values since the diameter of the vessel enlarged to the same imposed by the stent [50].

This procedure reduces patient discomfort and post - procedure complications, as it has the theoretic advantage of decreasing the risk of vessel recoil and recurrent stenosis [50].

3. Clinical Trials - Repeatability

It is important to evaluate the inherent precision of an equipment before proceeding into an advanced clinical validation. This chapter defines the term repeatability and focus on how to assess repeatability with the help of descriptive statistics, correlations techniques and the Bland – Altman method.

3.1 Repeatability

Repeatability can be understood as the variability of measurements obtained by one person while measuring the same item repeatedly. In other words, repeatability is the inherent precision of the measurement device [51]. Considering the two probability density functions (Figure 3.1), with two different measurements (A and B), the density functions demonstrate that measurement B is more repeatable than measurement A.

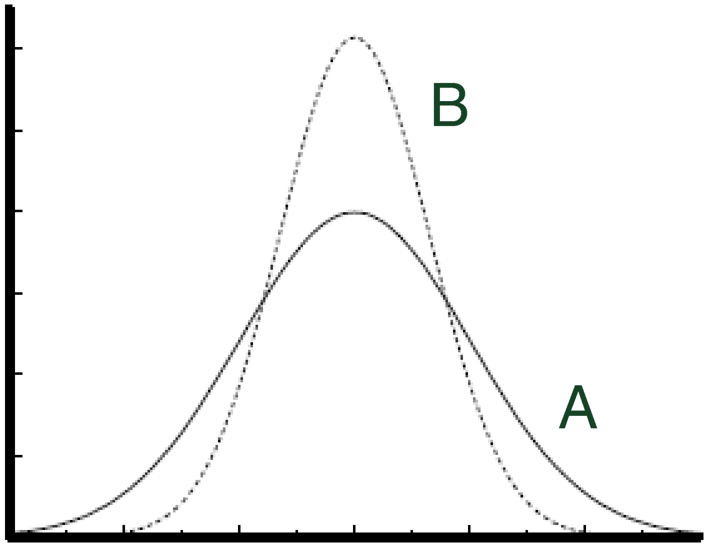


Figure 3.1 – Probability density function between two measurements (A and B). Adapted from [50];

The best way to examine and assess repeatability is to take repeated measurements on a series of subjects. According to the Guidelines for Evaluating and Expressing the Uncertainty of National Institute of Standards and Technology (NIST) Measurement Results [52], the following conditions need to be fulfilled in the establishment of repeatability:

- The same measurement procedure;
- The same observer;
- The same measuring instrument, used under the same conditions;
- The same location;
- Repetition over a short period of time;

3.2 Descriptive statistical analysis

Before assessing the repeatability of a specific equipment, it is important to know how is the data. And so, the first step in assessing repeatability is to “know” the data we have at our disposal using statistical methods.

3.2.1 Normality assessment

A normal distribution is used to describe a symmetrical, bell – shaped curve, which has the greatest frequency of scores in the middle, while the smaller frequencies lay in the extremes (Figure 3.2) [53, 54]. A non – normal distribution is visualized when the normal distribution definition isn’t fully respected. Usually, a non – normal distribution can be acknowledged when we have a non – symmetrical distribution, with higher extremes frequencies than in the usual normal distribution [53, 54].

Before any data analysis, it is important to check if the variables in the set of data have a normal or non – normal continuous distribution. This normality assessment of the variables in our dataset should always be considered, as some tests are risky for non – normal data. The best quantitative way to assess normality in continuous data is to apply the Kolmogorov – Smirnov (KS) one – sample test, where the maximum difference between the sample cumulative distribution and the hypothesized cumulative distribution are compared [55].

3.2.2 Central tendency measurement

The most common and effective central tendency numeric measure is the arithmetic mean, which is the sum of the values that compose variable divided by the size of the collection**.** However, the mean is not always the best way of measuring the center of the data, due to its sensitivity to outlier (e.g. extreme) values. Even a small number of extreme values can corrupt the mean. This effect can be counterbalanced by using the trimmed mean, which involves the calculation of the mean after chopping off a small percentage of the highest and smallest values. However, we should know that trimming too large portions (such as 20 percent) at both ends can result in a loss of important information [56].

When we are dealing with non – symmetric data, a better measure of the center of data is the median. In a dataset of N distinct values, sorted in numeric order, if N is odd, then the median is the middle value of the numeric ordered set. If N is even, the median is the average of the middle two values [56].

3.2.3 Data dispersion measurement

The degree to which numeric data tends to spread is the dispersion, or variance of the data. Standard deviation (SD) is the most common dispersion measure and shows how much variation exists from the average. A low SD indicates that the data points are very close to the mean, while high SD values indicates that the data points are spread out over a large range of values [56].

The k^{th} percentile of a set of data in numeric order is the value x_i having the property that k percent of the data entries lies at or below x_i. Quartiles are another type of commonly used percentiles. The first quartile (Q_1) is the 25th percentile; the third quartile (Q_3) is the 75th percentile. One must also perceive that the median of a set of data is equivalent to the 50th percentile and second quartile (Q_2) [56]. The distance between the first and third quartiles is a simple measure of spread that gives the range covered by the middle half of the data. This distance is the interquartile range (IQR) and is defined as:

$$IQR = Q_3 - Q_1 \tag{3.1}$$

A common rule of thumb for identifying suspected outliers is to single out values falling at least 1.5 × IQR below the first quartile or above the third quartile, while extreme values fall at least 3 × IQR below the first quartile or above the third quartile.

The full quantitative summary of the shape of a distribution is often known as the five – number summary of a distribution, which consists of the median, Q_1, Q_3 and the minimum and maximum values of the distribution. Those quantitative measure are then, placed in order: minimum, Q_1, median, Q_3, maximum [53, 56].

3.2.3.1 Boxplot

The five – number summary can also be expressed graphically by a boxplot (also known as box-and-whisker diagram), which is one of the most popular ways of visualizing distributions and comparing sets of compatible data (Figure 3.2). The ends of the box are at the quartiles, so that the box length is the IQR. The median is marked by a line within the box, and the superior and inferior lines are the whiskers. The whiskers are extended to the extreme low and high observations only if these values are less than 1.5 × IQR beyond the quartiles. Otherwise, the whiskers end at the most extreme observations occurring within 1.5 × IQR of the quartiles and every value outside that range is considered as an outlier value and is represented with a dot. Values outside the 3 × IQR are considered as extreme values, and are represented with an ‘x’ mark on the boxplot figure [56].

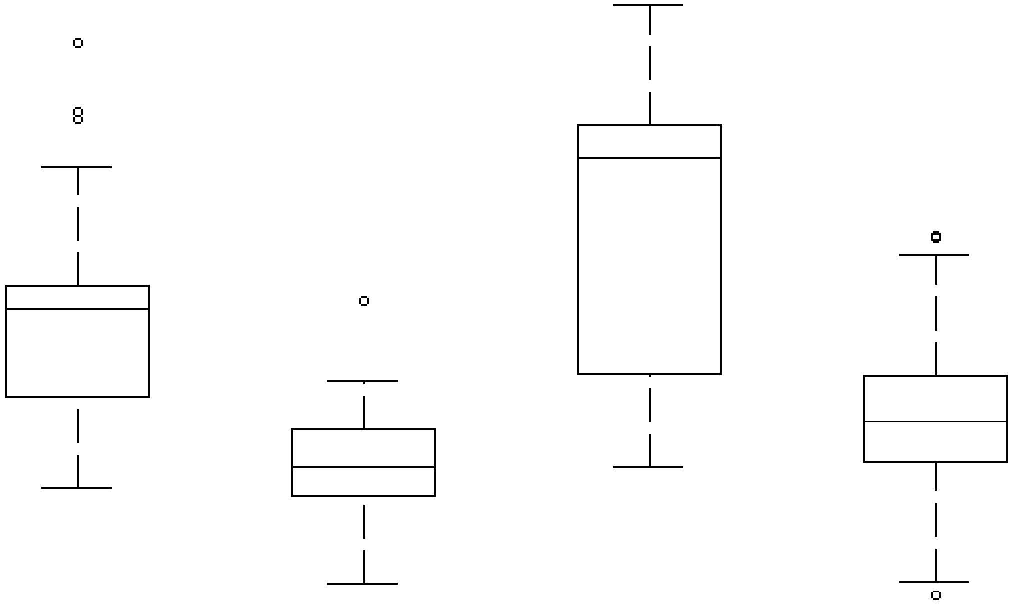


Figure 3.2 – Representation of four boxplots with different dispersion levels. Outliers can be visualized as the dots outside the whiskers;

3.3 Statistical tests for group comparison

There is a whole family of statistical techniques that can be used to test for significant differences between groups, consequently testing different hypotheses. The most important parametric and non – parametric techniques for group comparison will be briefly covered on this section, with the main focus on how and when to use each technique. The techniques will not be described with further statistical detail, as it is beyond the scope of this dissertation.

3.3.1 Null hypothesis

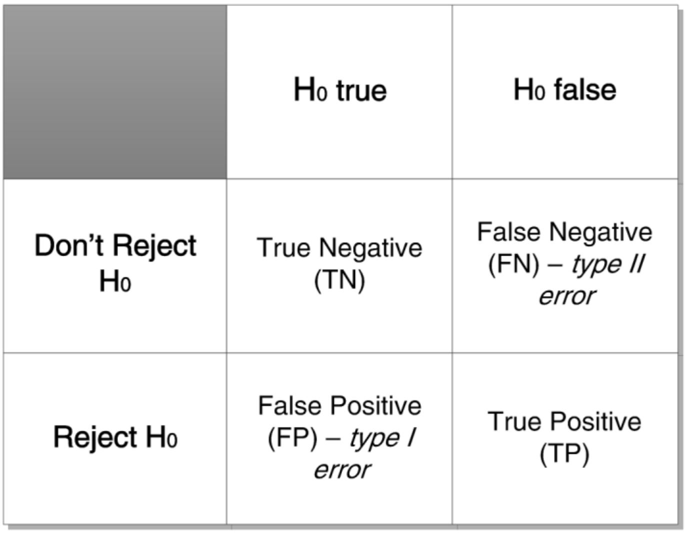
The structure of hypothesis testing is formulated with the use of the term null hypothesis (H_0) referring to any hypothesis that is wished to be tested. Usually, the analyst arrives at one of two possible conclusions: accept H_0 or reject H_0. The rejection of H_0 leads to the acceptance of an alternative hypothesis (H_1) [54].

H_0 is accepted or rejected based on the significance level (α), which is often 0.05, or 0.01. The null hypothesis H_0 is rejected when the p – value (estimated probability of rejecting H_0 when the hypothesis is true) is lower than α. When the null hypothesis is rejected, the result is said to be statistically significant [53, 54].

3.3.2 Type I and type II errors

As the purpose of parametric and non – parametric tests is to test hypothesis, there is always the possibility of reaching the wrong conclusion.

There are two types of errors that one can make. We may reject the null hypothesis when it is, in fact, true. In other words, we accept the hypothesis that there is a difference between our groups, but there is not. This is the type I error (or false positive (FP)). Although, we can also fail to reject the null hypothesis when it is, in fact, false (when the groups do not differ, when in fact they do), therefore occurring the type II error (or false negative (FN)). These two errors are inversely related. As we try to control for a type I error, we actually increase the likelihood of committing a type II error. One should always consider the power of a test and the sample size when evaluating the null hypothesis. We can also minimize the possibility of errors by choosing an appropriate p – value (usually .05/.01) [53, 55]. Figure 3.3 expresses the four possible outcomes of a hypothesis test.



	H_0 true	**H_0 false**
Don't Reject H_0	True Negative (TN)	False Negative (FN) – *type II error*
Reject H_0	False Positive (FP) – *type I error*	True Positive (TP)

Figure 3.3 – Representation of the possible outcomes that can occur while accepting or rejecting H_0. Adapted from [54];

3.3.3 Parametric tests

Parametric methods can produce more accurate and precise estimates, having high statistical power. However, before applying any parametric tests, one should have to consider some general assumptions.

3.3.3.1 General assumptions

General assumptions apply to all of the parametric techniques and should always be checked before applying any parametric test to avoid misleading results. One should know that even if the probability of having misleading results with a parametric test is higher with the violation of one or more of the following rules, it doesn't exactly mean that the parametric test result is misleading [53]:

- **Normal distribution –** As discussed in 3.2.1, it is important to assess data normality as some parametric tests are not usually applicable for non – normal data. Fortunately, most of the parametric techniques are reasonably robust, and so, with large sample sizes, the violation of this assumption doesn't cause any major problem.
- **Continuous scale –** Parametric approaches assume that the dependent variable is measured at the interval or ratio level, instead of using discrete categories.
- **Independence of observations –** Each observation or measurement must not be influenced by any other observation or measurement. A violation of this assumption can be very serious, according to Stevens (1996) [57].
- **Random sampling –** The parametric techniques assume that the scores are obtained using a random sample from the population. This is often not the case in real – life research.
- **Homogeneity of variance –** Parametric techniques make the assumption that samples are obtained from populations of equal variances. The Levene's test for equality of variances must take a part in t-test and analysis of variance (ANOVA) techniques, even though these techniques are robust to violations of the homogeneity of variance.

If the researcher is insecure or if one or more general assumptions were violated, the use of a non – parametric test must always be considered (more on that on sub – chapter 3.3.4).

3.3.3.2 Independent-samples t-test

The independent-samples t-test is used when you want to compare the mean scores of a continuous variable in two different groups of people or in two different conditions [53]. This test will reveal if there is a statistically significant difference in the mean scores for two different groups.

3.3.3.3 One – way analysis of variance (ANOVA)

ANOVA is commonly used when we are interested in comparing the mean scores of more than two groups. It compares the variability in scores between two different groups with the variability within each of the groups. An F ratio is calculated which represents the variance between the groups, divided by the variance within the groups. A large F ratio indicates that there is more variability between the groups (caused by the independent variable) than there is within each group (referred to as the error term) [53].

There are two different types of one – way ANOVA [53]:

- **Between – groups ANOVA –** used to compare different subjects or cases in each of the groups.
- **Repeated – measures ANOVA** – used to compare differences between same subjects under different conditions, or measures at different points in time. Data needs to be paired for repeated – measures ANOVA to be used.

3.3.4 Non – parametric tests

As referred in the last sub – chapters, one should always be careful if any general assumptions were transgressed. If one or more general assumptions are violated, the parametric test results can be misleading, and therefore, a non – parametric approach should be considered. [53] Non – parametric tests don't have such strict requirements, and don't make assumptions about the underlying population distribution. However, they are less robust and sensitive, and may therefore fail to detect differences between groups that actually exist (or vice – versa) [58].

Non – parametric techniques are also useful for data measured on nominal and ordinal scales and for very small samples [53].

3.3.4.1 Non – parametric alternatives to parametric tests

Each parametric test has its non - parametric alternative, which is usually less robust. As described in 3.3.3, if one or more general assumptions are violated, one must consider the use of the non – parametric alternative that matches the unusable parametric test. In Table 3.1 you will find the parametric tests with the corresponding non – parametric surrogate.

Table 3.1 – Comparison between each parametric tests and their respective non – parametric alternative. Adapted from [53];

Parametric test	*Non – parametric alternative*
Independent samples t – test	Mann – Whitney U test
One – way between groups ANOVA	Kruskal – Wallis H test

3.3.4.2 Mann – Whitney U test

This technique is the non – parametric equivalent to the independent samples t-test and is used to test for differences between two independent groups on a continuous variable. However, instead of comparing the means, the Mann – Whitney U test compares the medians. It then converts the scores on the continuous variable to ranks across the two groups. It then evaluates whether the ranks for the two groups differ significantly, or not. As the scores are converted to ranks, the actual distribution of the scores is irrelevant [53].

3.3.4.3 Kruskal – Wallis H test

The Kruskal – Wallis H test is the non – parametric alternative to the one – way between groups ANOVA. It is similar to the Mann – Whitney U test (see sub-chapter 3.3.4.2), but it allows the user to compare more than just two groups. Just as Mann – Whitney test, Kruskal – Wallis H compares the medians, and not the means. Scores are converted to ranks the mean rank for each group is compared. As this is a 'between-groups' analysis, different people must be in each of the different groups [53].

3.4 Correlation analysis

Correlation analysis is used to describe the strength and direction of the linear relationship between two variables [53]. Correlation coefficients provide a numerical summary of the direction and the strength of the linear relationship between two variables. The relationship between variables can be inspected visually by generating and inspecting a scatterplot, which will provide information on both the direction of the relationship (positive or negative) and the strength of the association.

One should know that correlation coefficients indicate linear relationship between variables. However, two variables can be related in non – linear fashion (e. g. curvilinear), and so, one should always check the scatterplot, especially if low values of *r* are obtained. Also, outliers can have a dramatic effect on the correlation coefficient, especially in small samples. One should careful analyze the scatterplot to check for outlier and/or extreme values.

Two main tests are used to assess correlation: the parametric Pearson product – moment correlation (*r*) and the non – parametric alternative Spearman rank order correlation.

3.4.1 Pearson product – moment correlation

The Pearson's correlation test is a parametric test which is designed for detecting linear relationships in continuous variables. It can also be used if you have one continuous variable and one nominal variable. Pearson correlation coefficients (*r*) can range from -1 to +1. The sign out the front indicates whether there is a positive correlation (as one variable increases, so does the other) or a negative correlation (when one variable increases, the other decreases). The size of the absolute value (while ignoring the sign) provides an indication of the strength of the relationship. A perfect correlation of -1 or -1 indicates that the value of one variable can be determined exactly by knowing the value on the other variable. On the other hand, a correlation of 0 indicates no relationship between the two studied variables [53].

However, the output can range between -1.00 and 1.00, so one should know how to interpret these values, and when should we consider small, medium and high relationship between two variables. Cohen (1988) [59] has suggested the guidelines presented in Table 3.2.

Table 3.2 – Guidelines for correlation level based on the value of the correlation coefficient range, according to Cohen *et al.* (1988) [59];

Correlation coefficient range	*Correlation*
r = 0.10 to 0.29 or r = -0.10 to 0-.29	Small
r = 0.30 to 0.49 or r = -0.30 to -0.49	Medium
r = 0.50 to 1.00 or r = -0.50 to -1.00	Large

3.4.2 Spearman's rank order correlation

Spearman's rank order correlation is used to calculate the strength and direction of a monotonic association between two variables, without making any assumptions about the frequency distribution of the variables. This is the non-parametric alternative to Pearson's product-moment correlation, and should be used instead of Pearson's product moment correlation for highly skewed data or when monotonic relationships are suspected [53, 60]. However, some experts believe that Pearson's and Spearman's test could be used together, to assess by comparison if associations are mostly linear, or monotonic [60].

3.5 Bland – Altman method

Bland and Altman (1986) [16] have noted that the use of correlation is quite misleading, because and high correlation does not necessarily mean that two methods/repeated measurements agree. A correlation coefficient only measures the strength of a relation between two variables, not the agreement between them. Because of this, data that seems to produce high correlations can have poor agreement. We only have complete agreement between two measurements if all the points lie along a line of equality. Also, correlation highly depends on the range of the true quantity of the sample and of the group of subjects selected. Bland and Altman also concluded that a lack of agreement between different measurements is almost inevitable, but what matters is the amount by which the respective different measurements disagree [16].

Therefore, Bland and Altman have proposed an alternative approach to assess agreement in clinical measurement, by plotting the difference between two measurements against their mean, allowing the possibility to investigate any possible measurements errors and/or lack of agreement. One should also consider the limits of agreement, in which most of the differences are expected to lie. To construct a Bland – Altman plot, one should estimate the mean difference (*d*) and the SD of the differences (*s*). The limits of agreement must be adjusted by subtracting or adding *d* for each limit of agreement. And so, we would expect most of the differences to lie between d – 2s and d + 2s *(*or d – 1.96s and d + 1.96s, to be more exact). It is expected that the differences between both assessments lies within the limits of agreement with approximately 95% probability [16]. An example of the Bland – Altman plot is shown on figure 3.4.

When using the Bland – Altman plot, both measurements should be paired, and independent from one another, to avoid inaccurate results. The mean difference (also known as bias) should be zero or as closest to zero as possible. If the bias is significantly different from zero, there is worrying discrepancies between methods or measurements, due to problems with the process of measurement or if both measurements are not independent from one another [16]. It is also important to analyze the variability of the plot and to pay attention for unusual patterns in a Bland – Altman plot [16].

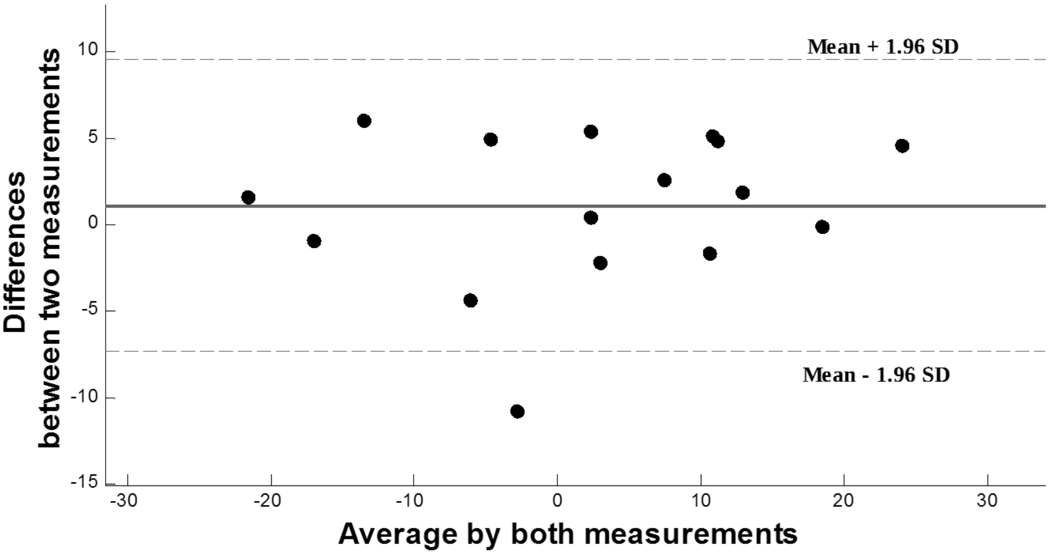


Figure 3.4 – Example of a Bland – Altman plot;

3.6 Related work

Repeatability measurements regarding the usefulness non – invasive techniques for CV function assessment is a commonly discussed issue in the scientific community. Among several studies, Crilly *et al.* (2007) [15] studied the repeatability of a radial applanation tonometer for pulse wave analysis (PWA). They proved that PWA demonstrates high levels of repeatability even when used by relatively inexperienced staff and has the short - term potential to be included in clinical practice. Frimodt – Møller *et al.* (2007) [61] assessed mean day – by – day differences in AIx in patients with chronic kidney disease, with AIx = 2.6 ± 11.2.

Results of validations with the purpose of evaluating the repeatability of new devices have also been reported. One of these new devices is Arteriograph (TensioMed, Budapest, Hungary), an oscillometric method. Horváth *et al.* (2010) [62] validated Arteriograph by comparing aortic AIx, SBP and PWV measurements during cardiac catheterization with the values measured by the Arteriograph. This experiment has shown that AIx, SBP and PWV oscillometric measures showed strong correlation with the invasively obtained values.

4. Data Mining

Data mining has attracted a great deal of attention in the information industry and in society as a whole in recent years. Firstly, this chapter introduces the concept and describes all the data pre – processing routines needed to efficiently execute data mining techniques. Then, special focus is given to classification and clustering techniques, which were subsequently used to extract new information from a database.

4.1 Data mining – Introduction

Being considered as the most important step in knowledge discovery from data (KDD), data mining is usually defined as the automatic or semi – automatic process of discovering patterns in data. When data mining techniques are efficiently applied, it can be possible to extract implicit, previously unknown and potentially useful information from data [63]. Due to this, data mining techniques have attracted a great deal of attention and recognition since the start of the new century, due to wide availability of huge amounts of data and the imminent need for turning such data into knowledge [64].

Nowadays, in industry, media and in database research communities, and adopting a broad view of data mining functionality, data mining is acceptably treated as a synonym for KDD [45]. A full view of the data mining process can be observed on figure 4.1.

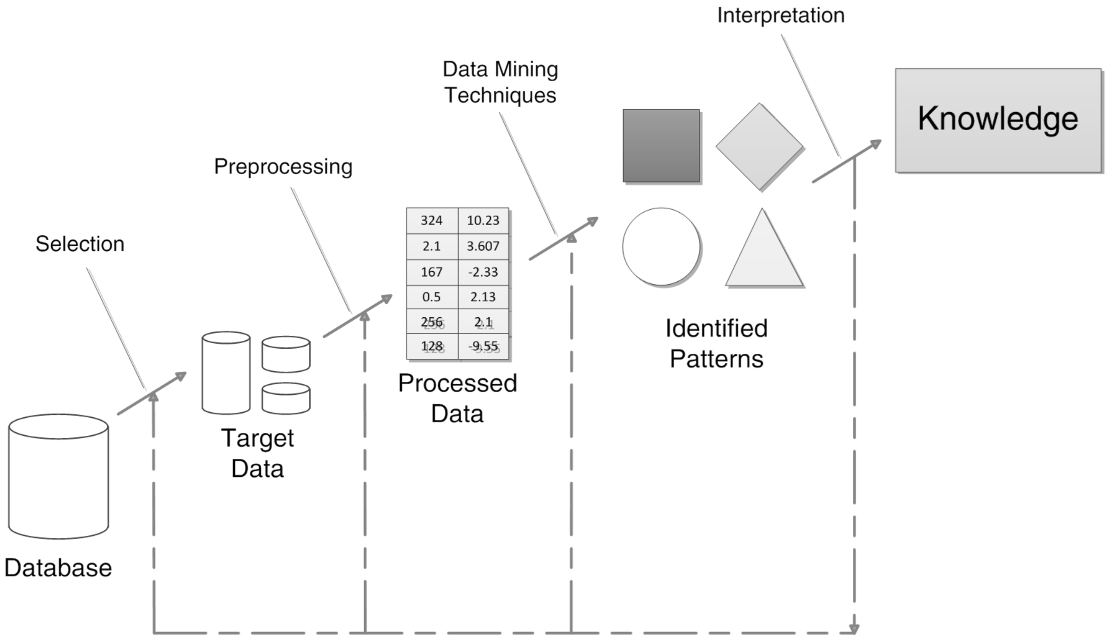


Figure 4.1 – Steps of KDD, commonly designated in new century as data mining. Adapted from [64];

4.1.1 Differences between statistics and data mining

Although they may seem quite similar, statistics and data mining procedures have their differences. First, statistics tend to use conservative analysis strategies based on rigorous mathematical approaches. Although data mining methods are based on mathematics, many techniques adopt a heuristic approach to solve real – world problems. Secondly, while statistical analysis is deductive, data mining is inductive [65]. In statistics, a hypothesis is built and then data is collected to test the hypothesis, as modern science does. In other words, statistics is a process of reasoning from the general to the specific. Data mining can work without a hypothesis, as it explores data that have been collected in advance and discovers hidden patterns from it. And so, data mining is process of producing general from the specific [66].

4.2 Data pre – processing

Usually, and before applying any statistical data mining techniques, the data we wish to analyze is incomplete, noisy, and inconsistent. As low – quality data will usually lead to low – quality mining results, data preprocessing becomes a fundamental step in the data mining process. Data pre - processing techniques, when correctly applied, can substantially improve the accuracy and efficiency of data mining techniques [64].

4.2.1 Descriptive statistics analysis

A preliminary analysis using descriptive data techniques is essential not only to know and understand the true properties of the data we wish to mine, but also to identify noise and outliers in the dataset. Normality, central tendency and data dispersion measurements are generally used to characterize data. These measurements may be expressed analytically or graphically. All the descriptive statistical methods were previously described in section 3.2.

4.2.2 Data cleaning

Data cleaning processes intend to discover and correct discrepancies in the data, which may be caused by diverse factors (i.e. human error, errors in instrumentation devices that acquire data, erratic code). A preliminary descriptive analysis of the data is essential to discover these inconsistencies. Usual data cleaning routines involve missing values removal, noise smoothing and outlier identification [64].

4.2.3 Irrelevancy and redundancy analysis

In a raw database, attributes that are redundant or irrelevant may slowdown and display unsatisfactory results. Relevance and redundancy analysis should be performed to detect and remove attributes that do not contribute to the data mining tasks.

For redundancy analysis, correlation tests can be used to identify whether any two given attributes are statistically related. For example, a strong correlation between two specific attributes may suggest that one of the two could be removed from further analysis. As for irrelevancy, attribute subset selection (or feature subset selection) can be used to identify and remove as many irrelevant and redundant features as possible. Irrelevancy and redundancy tests help in reducing the dimensionality of the data and will enable data mining techniques to operate faster and more effectively [64, 67].

4.2.4 Data normalization

Most of the times, data needs to be consolidated into forms appropriate for mining. One of the most common transformations is data normalization, which involves scaling all values for a given attribute so that they fall within a small specified range, such as -1.0 to 1.0 or 0.0 to 1.0. Usually, normalization procedures are necessary when attributes with initially large ranges outweigh attributes with smaller ranges (for example, in ANN) [64]. The most common data normalization method is min – max normalization. Considering that a value V from an attribute A with a minimum and maximum value (min_A and max_A, respectively) should fit in the range [C, D]. Then the normalized value (V') can be computed by the following formula:

$$V' = \frac{(V - min_A)}{(max_A - min_A)}(D - C) + C \tag{4.1}$$

4.2.5 Data discretization

Data discretization techniques can be used to reduce the number of values for a given continuous attribute by dividing the range of the attribute into intervals. Interval labels can then be used to replace actual data values. Usually, replacing numerous continuous values by a small number of interval labels reduces and simplifies the original data, consequently leading to a concise knowledge – level representation of data mining results [64]. On the other hand, many techniques don’t deal well with discretized attributes [67].

4.3 Data classification

Data classification is the process of constructing a model that describes and distinguishes data classes or concepts for the purpose of being able to use the classifier to predict the class of objects whose class label is unknown (Figure 4.2). Because the class label of each training tuple is given, data classification is an example of a supervised learning process [63, 66].

While Figure 4.2 represents the many steps of data classification, the process can be simplified as a two – step process,. The first step is the training step, where a classification model is built by analyzing pre – existent class labeled data. The second step, the testing step, examines a classifier for its accuracy using testing data. Subsequently, the model can be used for classifying objects with unknown label [66]. Sometimes, another independent set comes to surface: the validation set, which is used before the test set to minimize the error.

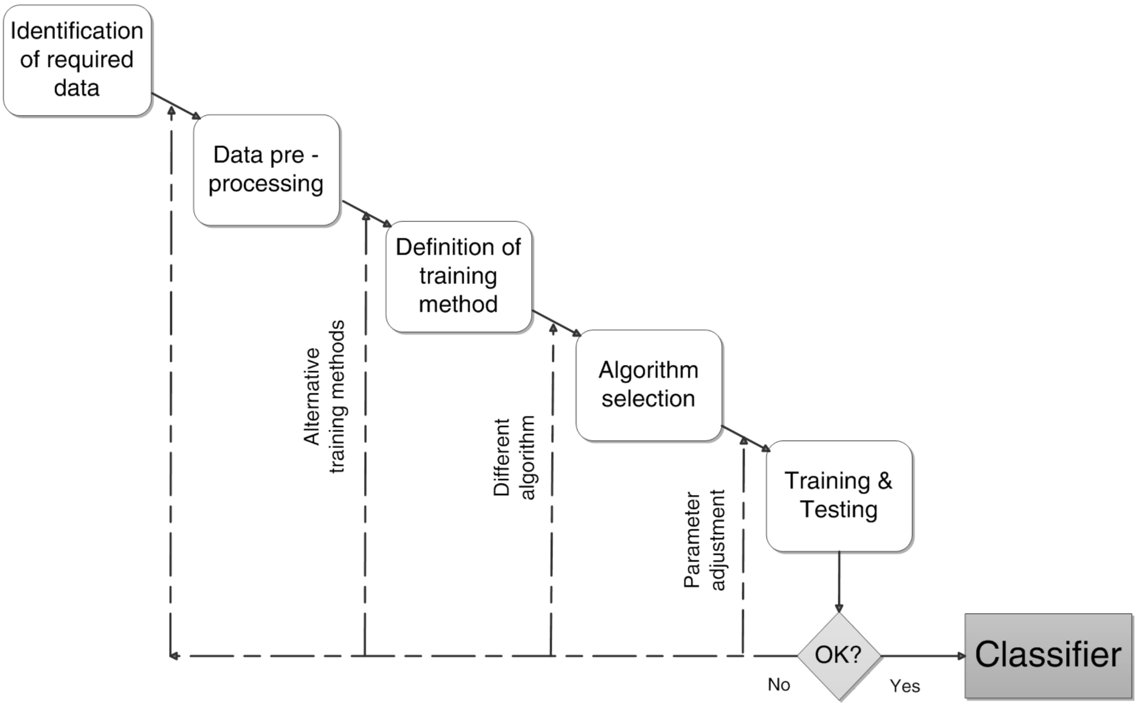


Figure 4.2 – Diagram representation of the steps of data classification;

4.3.1 Classification techniques requirements

Classifiers can be evaluated and compared according to the following criteria [64]:

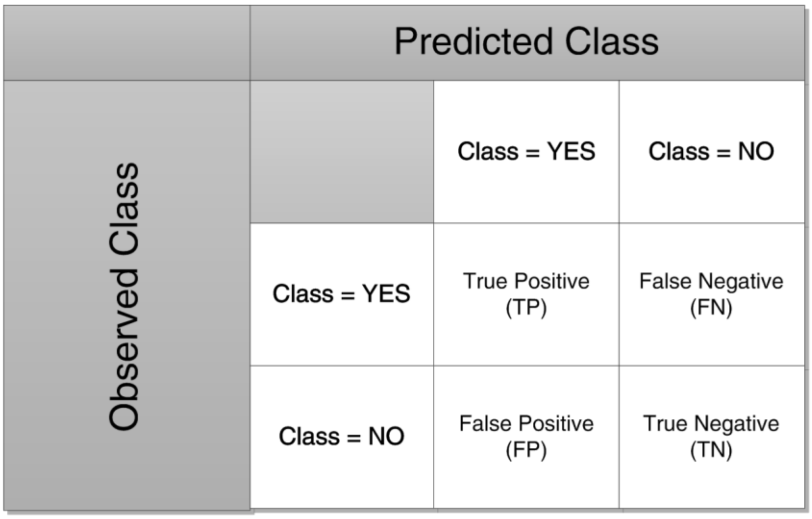
- **Accuracy** – The accuracy of a classifier refers to the ability of a given classifier to correctly predict the class label of new data.
- **Speed -** Speed indirectly refers to the computational costs involved in generating (learning speed) and using (classification speed) the classifier.
- **Robustness/Tolerance -** The robustness of a classification technique is defined by the ability of making correct predictions given incomplete and noisy data. Robustness also considers the algorithm dexterity in overcoming irrelevant and redundant attributes.
- **Versatility** – Versatility refers to the classifier's ability in dealing with different types of attributes (continuous/discrete/binary).

- **Comprehensibility -** Comprehensibility refers to the level of understanding and insight that is provided by the classifier. Comprehensibility is, however, very subjective and difficult to accurately assess.

4.3.1.1 Accuracy evaluation

The accuracy of a classifier refers to the ability of a given classifier to correctly predict the class label of new or previously unseen data. The accuracy of a classifier on a given test set is the percentage of test set tuples that are correctly classified by the classifier. We can also speak of the error rate or misclassification rate of the classifier, which is simply 1 – M, where M is the accuracy of the respective classifier. Although, we must keep in notice that the error rate of the model is optimistic of the true error rate, because the model is not tested on any samples that it has not already seen [63, 64]. Also, using training data to derive a classifier and estimate the accuracy of the resulting learned model can result in misleading overoptimistic estimates due to overspecialization of the learning algorithm to the data. So, accuracy is better measured on a class – labeled test set that wasn’t used to train the model.

As the performance of a classification process is based on the counts of testing objects correctly and incorrectly predicted, the confusion matrix (Figure 4.3) is a useful tool in analyzing how well your classifier can recognize objects of different classes. Given *m* classes, a confusion matrix is a table of at least size *m* by *m*. For a classifier to have good accuracy, ideally most of the tuples would be represented along the main diagonal of the confusion matrix, with the rest of the entries being close to zero [64].



		Predicted Class	
Observed Class		Class = YES	Class = NO
	Class = YES	True Positive (TP)	False Negative (FN)
	Class = NO	False Positive (FP)	True Negative (TN)

Figure 4.3 – Confusion matrix of observed class versus predicted class;

Given two classes, we can talk in terms of positive tuples versus negative tuples. True positive (TP) refer to the positive tuples that were correctly labeled by the classifier, while true negatives (TN) are the negative tuples that were correctly labeled by the classifier. FPs are the negative tuples that were incorrectly labeled as positives. Similarly, FNs are the positive tuples that were incorrectly labeled as negatives. And so, we can define accuracy as:

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{4.2}$$

However, accuracy measures may not always be enough. In a classifier to classifies medical data tuples as either healthy or hypertensive, a hypothetical accuracy rate of 90% may make the classifier seem quite accurate. But it could happen that only 3-4% of the tuples are actually "hypertensive", and therefore, an accuracy rate of 90% is not acceptable, because the classifier could be correctly labeling only the "healthy" tuples. In this case, sensitivity and specificity measures can and should be adopted. Sensitivity is also referred to as the TP positive rate, that is, the proportion of positive tuples that are correctly identified, while specificity is the TN rate, that is, the proportion of negative tuples that are correctly identified. Also, we may use precision to identify the percentage of tuples labeled as "hypertensive" that actually "hypertensive" tuples [64]:

$$Sensitivity = \frac{(TP)}{(TP + FN)} \tag{4.3}$$

$$Specificity = \frac{(TN)}{(TN + FP)} \tag{4.4}$$

$$Precision = \frac{(TP)}{(TP + FP)} \tag{4.5}$$

4.3.2 Definition of training set

After all the data pre – processing are performed, it is of uttermost importance the definition of how the training and testing will proceed, to obtain a reliable estimate of classifier accuracy, with methods like holdout, random subsampling and cross – validation among the most used.

4.3.2.1 Holdout method

In this method, the given data are randomly partitioned into two independent sets: a training set and a test set. Typically, approximately two – thirds of the data are allocated to the training set, and the remaining one – third is allocated to the test (70 - 30 is another usual ratio for dividing training and testing tuples). The training set is used to derive the model, whose accuracy is estimated with the test set. However, this estimate is considered as pessimistic, because only a portion of the initial data is used to derive the model [64].

Random subsampling is a variation of the holdout method, in which the holdout method is repeated *k* times. The overall accuracy estimate is taken as the average of the accuracies obtained [64].

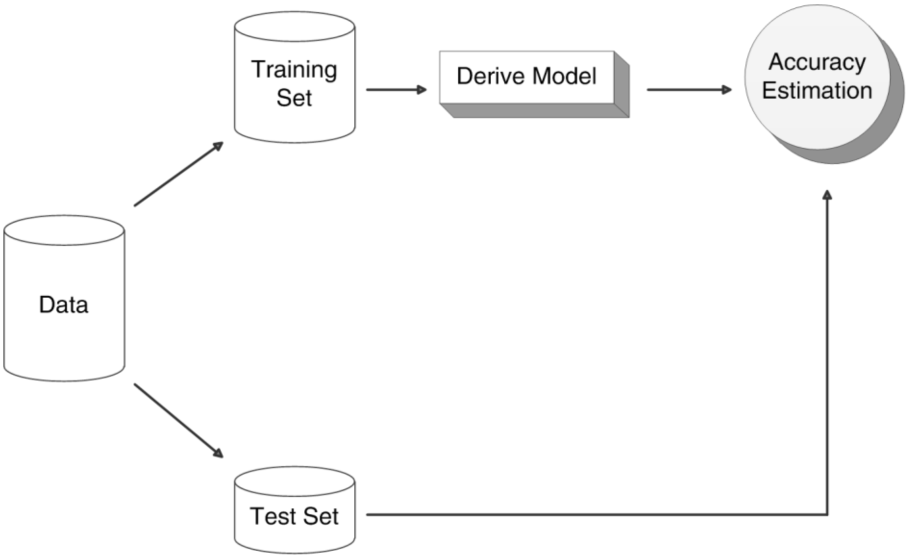


Figure 4.4 – The holdout method. Adapted from [64];

4.3.2.2 Cross - validation

In *k*-fold cross – validation (CVN), the initial data are randomly partitioned into k mutually exclusive subsets or "folds", D_1, D_2, …, D_k, each of approximately equal size. Training and testing is performed *k* times. In iteration I_1 partition D_1 is used as a test set, and the remaining partitions are collectively used to train the model. That is, in the first iteration, subsets D_2,…,D_k collectively serve as the training set in order to obtain a first model, which is tested on D_1; the second iteration is trained on subsets D_1,D_3,…,D_k and tested on D_2, and so on. Unlike the holdout and random subsampling methods above, here, each sample is used the same number of times for training and once for testing. The accuracy estimate is the overall number of correct classifications from the k iterations, divided by the total number of tuples in the initial data [64].

A more robust variation is stratified *k*-fold CVN, where the folds are stratified so that the class distribution of the tuples in each fold remains approximately equal to the initial data, minimizing the bias and variance in a dataset. Because of that, stratified 10 – fold CVN is highly recommended for estimating accuracy and should be used when possible. Otherwise, 10 – fold CVN is good enough [64].

4.3.3 Algorithm selection

After defining how the training and testing will be performed, we should select the algorithm that will be evaluated (however, this can be interchangeable in some cases). The algorithms that were used in this work will be further described on this sub – section.

4.3.3.1 Artificial neural networks (ANN)

Inspired by the biological nervous system, ANN are highly sophisticated analytical techniques, capable of modeling complex non – linear functions and predicting new observations from other observations. They appear as an effective and practical technology, and can be implemented by using electronic components or are simulated in specialized software on a digital computer [68].

Roughly speaking, an ANN is a set of connected input and output units in which each connection has a specific weight associated to it. During the learning phase, the network learns by adjusting the weights so as to be able to predict the correct class label of the input tuples [64]. The network is adjusted by comparing the output and the target, and the weights are adjusted until the network output matches the target (Figure 4.5). Usually, many input / target pairs are needed to successfully train a network [67]. Also, the network is feed – forward, which means that none of the weights cycles back to an input unit or to an output unit of a previous layer, making the information move in one direction only. Due to this, it is usual to see ANN often called as feed – forward neural networks or feed-forward ANN [64].

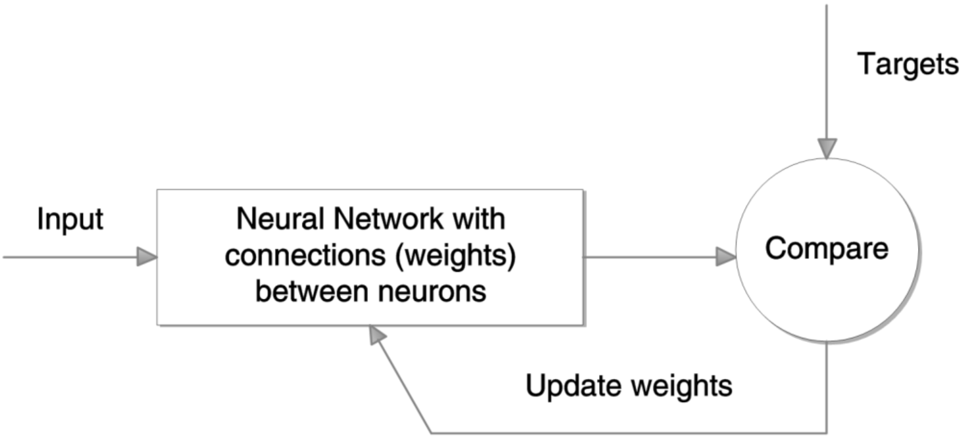


Figure 4.5 – ANN learning process. From [64];

ANN algorithms present important advantages that make their effectiveness undeniable, as their outstanding accuracy in general [64, 67] ability to work with continuous attributes and above average tolerance to redundant attributes [67]. However, they also have some drawbacks. Usually ANN need high computational power, and attributes need to be previously normalized. Also, NNs have poor interpretability, because they process information based on the 'blackbox' principle and, unlike other 'transparent' techniques, such as decision trees, they do not directly 'unveil' the way they process information [67].

4.3.3.1.1 Linear neural network (LNN)

Linear neural network (LNN; also known as single – layer perceptron (SLP)) is the simplest kind of feed – forward neural network. It is based on the McColluch and Pitts artificial neuron [69], which consists of a single input layer with input units and a layer of output nodes. The inputs to the network correspond to the attributes measured for each training tuple [64], and in this approach, the inputs are fed directly into the outputs. A LNN generic example can be found on Figure 4.6.

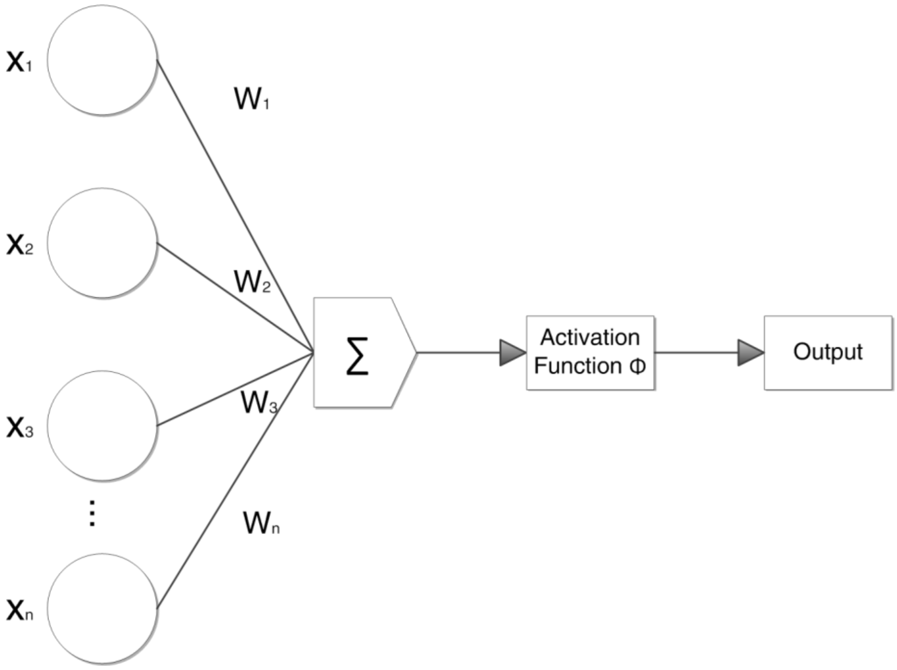


Figure 4.6 – A LNN generic example, with {X_1, X_2 … X_n} being input features and {w_1, w_2 … w_n} corresponding to connection weights. Adapted from [64];

If {X_1, X_2 … X_n} are input features and {w_1, w_2 … w_n} are connection weights (typically real numbers in the interval [-1, 1]), then the perceptron computes the sum of weighted inputs:

$$\sum_i x_i w_i \tag{4.6}$$

The output then goes through an activation function Φ, which is usually a threshold function. If the sum is above the desired threshold, output is 1; else, the output is 0. The most common method that the LNN algorithm learns from a batch of training instances is to run the algorithm repeatedly through the training set until it finds a set of connection weights which is correct on all training set instances. Afterwards, the trained model is used for predicting the labels on the test set [66].

Nowadays, LNN algorithms have been overwhelmed by other more sophisticated ANN. Still, they provide themselves as a good alternative and as a benchmark against which to compare the performance of more complex ANN [32].

4.3.3.1.2 Multi - layer perceptron (MLP)

Multi - layer perceptron (MLP) is the most popular ANN architecture in data classification techniques. This may be due to the conclusion that MLP with one or two hidden layer are universal approximators in a very precise sense [70, 71]. The MLP consists of a set of source units that constitute the input layer, one or more hidden layers of neurons and an output layer. From a statistical point of view, they perform nonlinear regression. An example of a multilayer feed – forward network is shown on Figure 4.7.

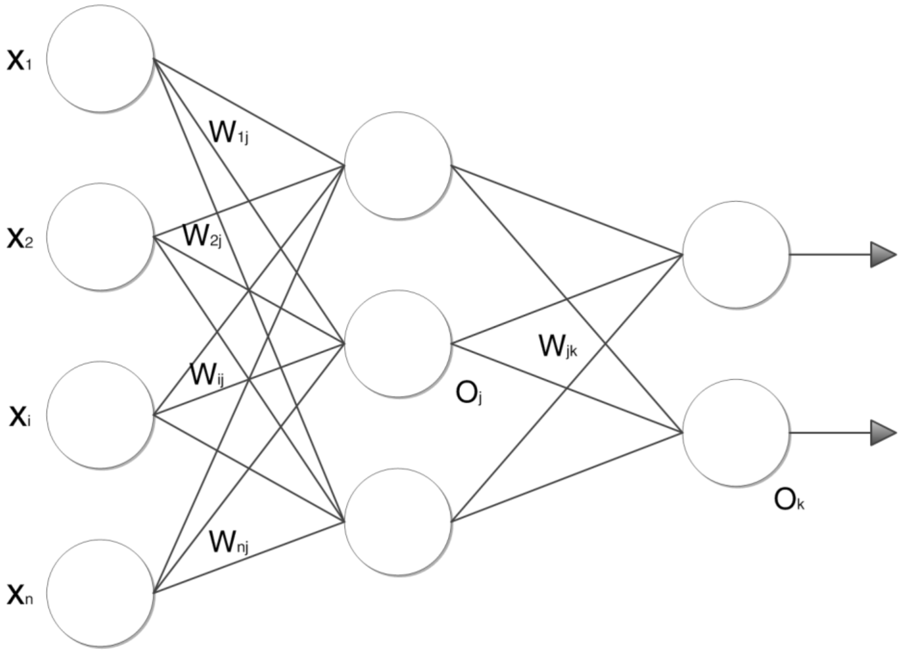


Figure 4.7 – A MLP generic example, with {X_1, X_2 … X_n} being input features and {w_{1j}, w_{2j} … w_{nj}} and W_{jk} corresponding to connection weights between the input layer and the next layer. O_j is the output computed by neuron *j*, and O_k is the output computed by neuron *k*. Adapted from [64];

Like the LNN, the inputs to the network correspond to the attributes measured for each training tuple, and are fed simultaneously into the input units, making up the input layer. After the network inputs pass through the input layer, they are weighted and fed simultaneously to a second layer known as the hidden layer, composed of hidden units (sometimes referred to as neurodes). The outputs of the hidden layer can be inputs to another hidden layer or inputs to output units that make up the output layer, which emits the network's prediction for given tuples [64]. An MLP with one input layer, one hidden layer and one output layer is characterized as a two – layer MLP. The input layer is not counted as it serves only to pass the input values to the next layer. Similarly, an ANN containing two hidden layers in called a three – layer MLP, and so on [64].

As a learning algorithm, the most well – known and widely used is the backpropagation algorithm, which learns by iterative processing of a data set of training tuples, comparing the network's prediction for each tuple with the known target class label. For each training tuple, the weights are updated so as to minimize the mean squared error between the network's prediction and the actual target value. These modifications are directed backwards, that is, from the output layer through each hidden layer down to the first hidden layer (hence the name backpropagation). In general, the weights will eventually converge and the learning process stops [64].

The main problem of the MLP is in defining the most correct topology, or finding the topology that gives us the best possible approximation for the problem we wish to model. Before training can begin, the user must decide on the network topology by perceiving the number of units in each layer. The number of input units and output units is determined by the number of initial attributes and known classes, respectively. Properly determining the optimal size of hidden units in the hidden(s) layer(s) is a challenging issue, because an underestimation can lead to poor approximation and generalization model capabilities [66], while an overestimation results in overfitting [64]. Also, there are no clear rules as a reference number of hidden units per hidden layer, and initial weight values also affect the model accuracy. The only solution is performance several trial-and-error, repeating the training process with different network topologies. Optionally, a different set of initial weights can also help in achieving the best possible MLP performance [64].

4.3.3.1.3 Radial basis function (RBF)

Radial basis function (RBF) represents an equally appealing and intuitive ANN, consisting of a hidden layer of radial units, in which every hidden unit implements a radial activation function and each output unit implements a weighted sum of hidden unit outputs. RBF training procedures is divided in two stages. In the first stage, the centers and widths of the hidden layer are determined by partitioning algorithms. Secondly, the weights connecting the hidden layer to the output layer are assessed by singular value decomposition or least mean squared algorithms [67].

RBF can be trained extremely quickly, much faster than MLP. On the other hand, the RBF is more sensitive to the curse of dimensionality [32]. Also, one should know that the danger of overfitting and underfitting is identical to MLP. Selecting the appropriate number of basis functions remains a critical issue and should not be ignored [67].

4.3.3.2 Decision trees induction

Decision trees induction is the learning of decision trees from class – labeled training tuples [64]. A decision tree assumes a flowchart structure, where each internal node (non-leaf node) in a decision tree represents a test in an attribute that needs to be classified, and each branch represents a possible outcome of the test. Each leaf node represents the outcome of the test. The topmost node is the root node, and instances are classified starting at the root node and sorted based on their feature values [67].

Decision trees induction usually adopts a nonbacktracking approach, and most of the decision trees are constructed in a supervised top – down recursive divide-and-conquer manner, starting with a training set of tuples and their associated class labels [64]. The training set is recursively partitioned into smaller subsets as the tree is being built. Afterwards, the constructed decision tree can be used as a classification algorithm, when a tuple with unknown class label is tested against the decision tree. A path is traced from the root to a leaf node, which holds the classification for the tuple.

Decision tree classifiers are quite popular, mainly because of their simplicity and comprehensibility. People can easily understand why a decision tree classifies and instance as belonging to a specific class [67]. Also, decision trees have great capacity of handling high dimensional data, and their induction is fast. Their construction does not require any domain knowledge or parameter testing and, therefore, is appropriate for exploratory data mining, having a good accuracy in overall. Decision trees tend to perform better when dealing with discrete/categorical features [67].

4.3.3.2.1 Random forest

Proposed by Breiman (2001) [72], random forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all the trees in the forest. The generalization error will eventually converge to a limit as the number of trees in the forest enlarges. After a large number of trees is generated, they vote for the most popular class [72].

4.3.3.2.2 C4.5

C4.5 is a very popular decision tree algorithm. Developed by Quinlan (1984) [73], it uses the decision tree induction approach of supervised top – down recursive divide-and-conquer tree construction. C4.5 is optimized for handling both continuous and discrete attributes and pruning the decision trees after their respective creation.

4.3.3.3 Bayesian classification

Bayesian classifiers predict class membership probabilities, such as the probability that a given tuple belongs to a particular class [64]. This type of classifiers is based on Bayes' theorem:

$$P(C|X) = \frac{P(X|C).P(C)}{P(X)} \tag{4.7}$$

Where P (C|X) denotes a posterior probability of an event C conditioned by an observation X and P (X|C) is the posterior probability of occurring observation X when event C is true. P(C) and P(X) are prior beliefs, where P(C) is the probability of the event occurring before observing any case X, and P(X) represents the probability of occurring a case X without considering any hypothesis C [19, 64].

The major advantage of the Bayesian classifiers is their short computational time for training. They also present great tolerance to missing values, and are an efficient approach in avoiding overfitting of data, because of their simplicity. However, because of this simplicity, their accuracy is often lower than other classifiers [66].

4.3.3.3.1 Naïve Bayesian classification

Let D be a training set of tuples with associated class labels. Each tuple is represented by an n-dimensional attribute vector X = {x_1,x_2,…x_n}, with the attributes represented as A_1,A_2,…A_n}. Considering a m-dimensional class vector C = {C_1,C_2,…C_m} and given a tuple, X, the classifier will predict that X belongs to the class having the highest posterior probability. The classifier predicts that X belongs to class C_i if and only if:

$$P\ (C_i|X) > P\ \left(C_j\middle|X\right) for\ 1\ \le j\ \le m, j\ \neq i \tag{4.8}$$

Thus, P(C_i|X) is maximized, and class C_i becomes the maximum posteriori hypothesis. By Bayes' theorem:

$$P(C_i\,|\,X) = \frac{P(X|C_i).P(C_i)}{P(X)} \tag{4.9}$$

Only P(X|C_i) needs to be maximized, because P(X) is constant for all classes. If the class prior probabilities are unknown, the classes are assumed to be equally likely, therefore, P(X|C_i) is maximized [63].

After that step, an assumption of class conditional independence is performed to reduce computation in P(X|C_i). This presumes that the attributes are conditionally independent of one another, given the class label of the tuple X [64]. In order to predict the class label of X, $P(X|C_i).P(C_i)$ is evaluated for each class C_i. The classifier will only predict that class label of tuple X is the class C_i if and only if:

$$P\ (X|C_i)P(C_i) > P\ \left(X\middle|C_j\right)P\left(C_j\right)\ for\ 1\ \leq j\ \leq m, j\ \neq i \tag{4.10}$$

In other words, the predicted class label is the class C_i for which $P(X|C_i)P(C_i)$ is the maximum [63, 64].

4.3.3.3.2 <u>Bayesian Networks</u>

The Naïve Bayesian classifier makes the assumption of the class conditional independence, that is, given the class label of a tuple, the values of the attributes are assumed to be conditionally independent of one another. While this simplifies computation, dependencies can exist between variables. Bayesian Networks provide a graphical model of causal relationships, specify joint conditional probability distributions that allow class conditional independencies to be defined between subsets of variables, and can be used for classification [64].

4.3.3.4 Associative classification

In associative classification, association rules that show strong associations between attributes-value pairs that occur frequently in a given dataset are generated. Typically, the discovery of association rules is based on frequent itemset mining. Associative classification algorithms are regarded as being easy to understand and having very good accuracy. However, they are really sensitive to noisy and incomplete datasets [67].

4.3.3.4.1 <u>RIPPER</u>

Proposed by Cohen (1995) [74] as an optimized version of the Incremental Reduced Error Pruning (IREP) [75], Repeated Incremental Pruning to Produce Error Reduction (RIPPER) is based on association rules with reduced error pruning, a very common and effective technique of decision tree algorithms.

In RIPPER, after a training set of data is split into a growing set and a pruning set, a single rule set is grown using one subset of the data and subsequently pruned. After the rule set has been generated, a simplification phase ensures that the pruning operator that yields the greatest reduction of error of the pruning set is chosen. Simplification ends when applying any pruning operator increases error of the pruning set [74].

4.3.4 Classification techniques comparison

As enunciated before, classifiers can be compared according to the criteria defined in in 4.3.1. Kotsiantis (2007) [67] has performed a review which compares the characteristics of each type of classifier, with * representing the worst and **** representing the best. This review was adapted to a synthetic table (Table 4.1). Only characteristics that were defined in the criteria exposed in 4.3.1 are identified and described in Table 4.1.

As exposed in Table 4.1, ANN present the highest accuracy of all the analyzed techniques and are highly versatile (even though they deal better with continuous attributes). However, they are quite sensitive to missing values and irrelevant attributes. Moreover, their training speed (with the exception of the RBF) and interpretability are the definitely worst among the four techniques. A rigorous pre – processing is absolutely needed when using ANN, to fully expose their potential [67].

As for other techniques, decision trees induction and associative classification also present good accuracy, and also excel in terms of versatility and comprehensibility, but decision trees are usually the best choice due to their intrinsic high tolerance. Naïve Bayes classification has the best training speed, high comprehensibility of the given results and high tolerance to missing values. But, on the other hand, it has the lowest accuracy of the comparison [67].

Table 4.1 – Comparison between the previously described classification techniques. * represents the worst result and **** the best result. Adapted from [67];

Classifier Characteristics		ANN	Decision Trees Induction	Naïve Bayes Classification	Associative Classification
Accuracy		****	***	*	***
Training Speed		*	***	****	**
Tolerance	**Missing values**	*	***	****	**
	Irrelevant attributes	*	***	**	**
	Redundant attributes	**	**	**	**
Versatility		***	****	***	***
Comprehensibility		*	****	****	****

4.4 Data clustering

Clustering is the process of grouping a set of objects into a class of similar objects (Figure 4.8). A cluster is a collection of data objects that are similar to one another within the same cluster, and are dissimilar to other objects contained in other clusters [64].

The clustering technique can be used in data mining processes, by partitioning the set of data into groups based on data similarity and then assigning labels to the number of groups created. Unlike classification, clustering is an example of unsupervised learning, which doesn't rely on predefined classes and class-labeled training examples, learning by observation and not by examples [64]. Clustering can also be used for outlier detection and as a preprocessing tool for other algorithms such as classification or attribute subset selection [63].

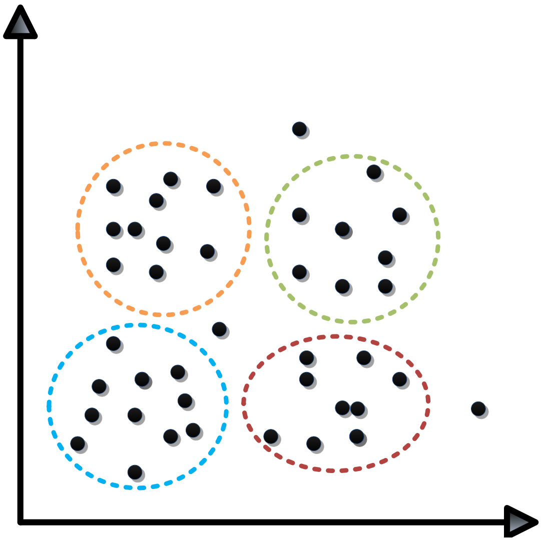


Figure 4.8 – The clustering process;

4.4.1 Clustering techniques requirements

Just like the classification techniques, clustering methods have their own special requirements, and can also be evaluated and compared, considering the following criteria [64]:

- **Interpretability –** The clustering results must be interpretable, comprehensible, and usable.
- **Scalability –** Highly scalable clustering algorithms are needed. Many clustering algorithms work well on small databases. However, large databases contain millions of objects, and clustering on a sample of given large data set may lead to biased results.
- **Dimensionality –** Clustering methods must be able to deal with high – dimension data. While many clustering algorithms are good at handling low –

dimensional data (few attributes), they give inconsistent results while dealing with high – dimensional datasets.

- **Robustness/Tolerance –** It is important that clustering methods have high noise and missing value tolerance.
- **Ability to deal with different types of attributes –** Clustering techniques must be able of dealing with different types of data instead of numerical and continuous data (binary, categorical, ordinal…).
- **Arbitrary shape detection –** It is important to develop algorithms that can detect arbitrary shapes. Many clustering algorithms use distance measures to determine clusters. Algorithms based on this type of measures tend to find spherical clusters with similar size and density, not considering that a cluster could be of any shape.

4.4.2 Cluster methods

Different types of clustering exist on the literature and can be classified into partitional, hierarchical, density – based, grid – based and model – based. In this work, mainly due to hardware and software limitations, only one partitional clustering approach and one model – based algorithm were used. And so, this dissertation will only focus on partitional and model – based approaches.

4.4.2.1 Partitional clustering

These algorithms decompose a database of *n* objects into a set of *k* partitions, where each partition represents a cluster and where each object must belong to exactly one group. The general criterion of a good partitioning is that objects in the same cluster are related to each other, whereas objects of different clusters are very different [76]. A popular approach of partitional clustering is the *k-means algorithm*, where each cluster is represented by the mean value of the objects in the cluster [64].

4.4.2.1.1 K – means algorithm

The *k*-means algorithm acknowledges the number of desired clusters input *k* and partitions a set of *n* objects into *k* clusters, so that the resulting intra - cluster similarity is high, but the inter - cluster similarity is low. Cluster similarity is measured in regard to the mean value of the objects contained in a cluster, which can be viewed as the cluster centroid [76].

The k-means algorithm starts with a random selection *k* of the objects, each of which initially represents a cluster mean. For each of the remaining objects, an object is assigned to the cluster to which it is the most similar, based on the distance between the object and the cluster mean. Afterwards, a new mean is computed for each cluster. This process iterates until the criterion function converges [64]. Conventionally, the square – error criterion based on the Euclidean distance function is used:

$$E = \sum_{i=1}^{k} \sum_{p \in C_i} |p - m_i|^2 \tag{4.11}$$

Where E is the sum of the square error for all objects in the data set, p is the point in space representing a given object; and m_i is the mean of cluster C_i (both p and m_i are multidimensional). In other words, for each object in each cluster, the distance from the object to its cluster is squared, and the distances are summed, resulting in k clusters that are quite compact and as separate as possible [64].

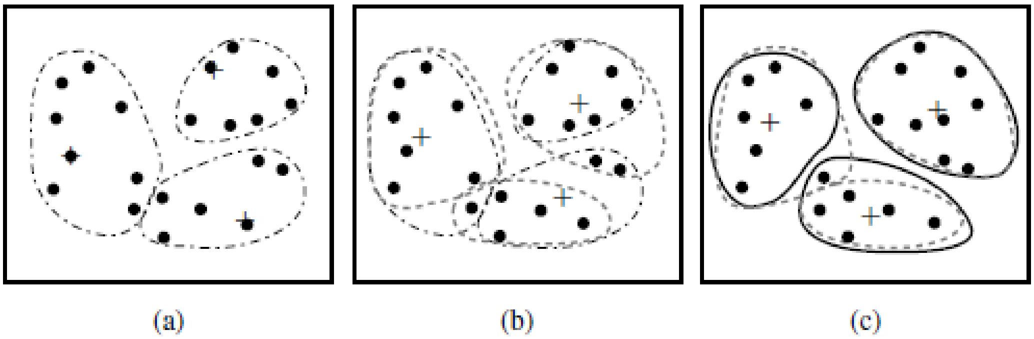


Figure 4.9 – K-means clustering with Euclidean distance function $k = 3$. A + represents the mean. of each cluster. From [64];

The k-means clustering method is relatively scalable and efficient in processing large datasets. However, it has some major disadvantages, as it is inapplicable to categorical attributes, it is unsuitable for discovering clusters with non - convex shapes of very different size and is quite sensitive to noise and outliers, because a small number of such data can substantially influence the mean value [64, 76].

4.4.2.2 Model - based clustering

Model – based methods hypothesize a model for each of the clusters and find the best fit of the data to the given model. It can also determine the number of clusters based on standard statistics, taking noise and outliers into account, yielding robust clustering methods. One well known approach is the Expectation – Maximization (EM) algorithm, which performs analysis based on statistical modeling [64].

4.4.2.2.1 Expectation – maximization (EM)

The EM algorithm is a popular clustering algorithm, which is, in fact, a complex probabilistic extension of the *k*-means method. Instead of assigning each object to the cluster with which it is most similar, EM assigns each object to a cluster according to a weight representing the probability of membership. This means that there are no strict boundaries between clusters and new means are computed based on weighted measures [64].

EM starts with an initial estimate of the parameters of the ‘parameter vector’, randomly selecting *k* objects to represent the cluster means (as in *k*-means partitioning). Then, EM iteratively refines the parameters (or clusters) based on an expectation step, where for each object x_i, the probability of cluster membership of each object for each of the *k* clusters, and on an maximization step, where we use the probability estimates from the expectation step to re-estimate the model parameters. These steps are performed until convergence is achieved [64].

The EM algorithm is a simple and easy alternative to the *k* – means approach, presenting good global performance in overall.

5. Hardware & Software

The developed prototype hardware and software systems are described in this chapter. The first sub-section briefly concentrates on the hardware, while the second sub-section focuses on the developed and integrated software, including acquisition and processing routines. The functionalities of the current MATLAB® software modules for acquisition and processing routines are also be presented.

5.1 Hardware

Hardware can be generically defined as the mechanical equipment necessary for conducting an activity. Hardware is a collective term, as it may include not only the computer, but also the cables, connectors, power supply units and other peripheral devices.

Our prototype hardware module for ADW acquisition consists of a PZ probe and signal conditioning circuit.

5.1.1 PZ sensor

A PZ element is able to convert force or pressure applied to its surface into a measurable voltage signal. PZ - based probes have been widely used in ADW measurements along the past few years due to their appealing characteristics: high sensitivity, high SNR and associated low – cost [13, 14]. The developed PZ probe is presented in Figure 5.1.

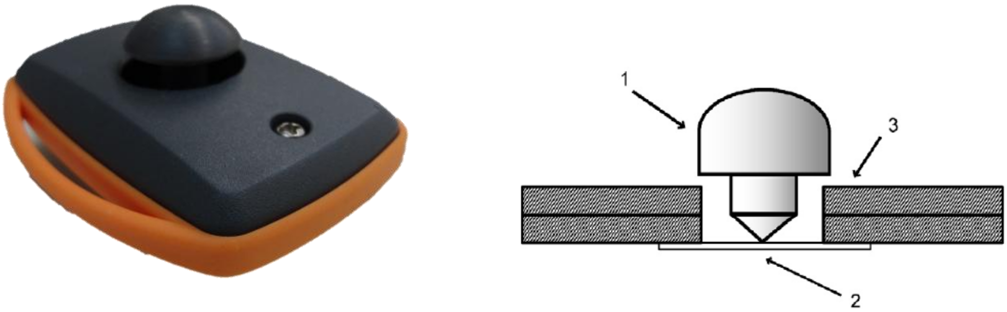


Figure 5.1 – PZ sensor. (a) outer upper – view, with the easily recognizable mushroom – shaped interface (b) probe elements in cut, with (1) mushroom – shaped interface, (2) PZ disc sensor and (3) printed circuit board (PCB). From [14];

The PZ probe consists of one PZ transducer bonded to a plastic block that supports bending under normal use. A mushroom polyvinyl chloride (PVC) interface was assembled over the PZ element. The pointy interface, which contacts directly with the sensor, exhibits the best performance in reproducing waveforms with low root mean square (RMSE) variance and, therefore, was the preferred option for ADW measurements [14].

5.1.2 Signal conditioning circuit

The ADW signal conditioning circuit architecture is shown on Figure 5.2, and consists of three main modules: a power supply module, a first amplifying stage and a processing module. The power is supplied via an USB cable that's connected to a personal computer. The computer supplies the needed voltage to the other two modules. In the first amplifying stage, the raw signal obtained from the PZ probe is amplified with a gain of 1000, using an active differentiator mode amplifier, proposed by Almeida et al. (2011) [14]. The signal then proceeds to the processing module, where a peak detector coupled with a timer is used to extract the reference time signal associated with the signal peak [14]. Since the signal is a time derivative of the physiological signal (due to sensor specifications), it is integrated by a Microchip® (Chandler, Arizona, USA) dsPIC33 microcontroller module.

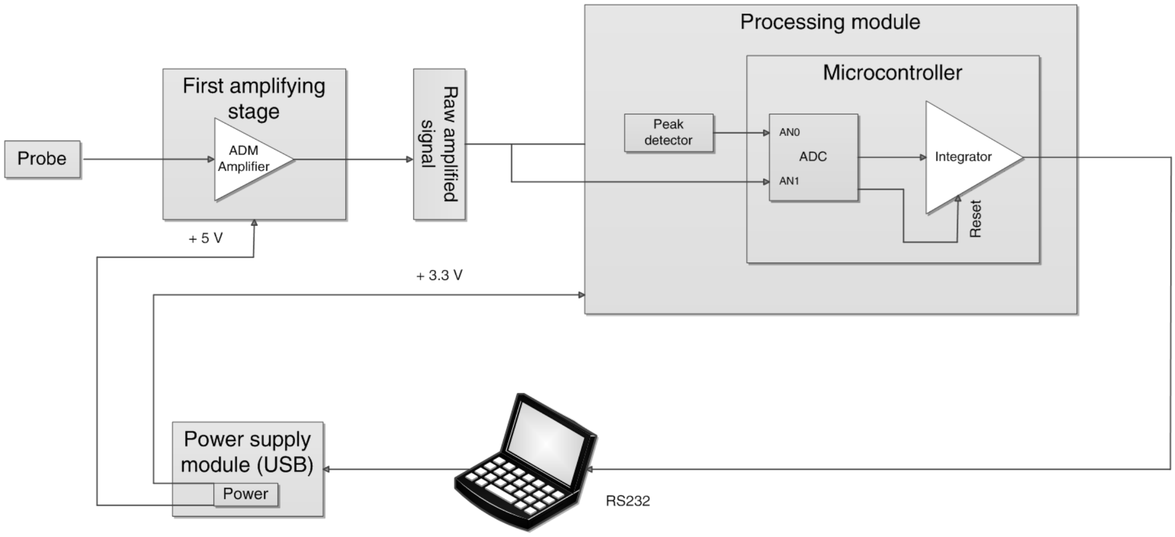


Figure 5.2 – Signal conditioning circuit. Adapted from [62];

5.2 Software

Software can be defined as the set of programs, procedures and algorithms related with the operation of a data processing system. In contrast with hardware, software is not touchable.

The algorithms for ADW acquisition and ADW processing modules were previously developed and integrated with the MATLAB® language (version 2009a). User – friendly GUIs were previously built for each module, resulting in an uncomplicated and user - friendly software arrangement. Several improvements were made in both of them during this project, which resulted in a clear time - saving optimization while adding some new and necessary features.

5.2.1 ADW Acquisition Module

The ADW Acquisition Module efficiently acquires the pulses from the subject when coupled with the previously described hardware. No major modifications were done, except one new button, which was added so that the ADW Acquisition Module could perform a direction communication with ADW Pulse Analyzer (more on that in sub – chapter 7.1.1).

5.2.2 ADW Pulse Analyzer

The acquired raw ADW may contain noise, artifacts and irregular waveforms. Also, each acquired pulse isn't segmented, and so, before applying any operation to retrieve information from the ADW, we must prepare the ADW data for feature extraction. The ADW Pulse Analyzer was developed to possess all the routines that are necessary to extract, from the acquired pulses, the features points that are necessary for a sequential statistical and data mining analysis with the help of baseline noise removal and pulse segmentation routines.

During this work, not only an aesthetic visual optimization was realized, but new functionalities were also implemented, including direct connections with the ADW Acquisition Module (see sub – chapter 7.1.1) and with an important stand – alone Bland – Altman GUI for repeatability assessment (see section 6.4.1.4.1). Other functionalities include processing suites for two different types of files: *.mat* (MATLAB® compatible) and *.arff* (Weka® compatible – see sub-section 6.4.2 and 6.4.3), a tool for automatic file merging and *.bmp* picture saving.

Most of the ADW processing routines were already developed and discussed in previous works [14, 24]. They will only be briefly presented in the next sub – chapters.

5.2.2.1 ADW onset calculation

The used algorithm for ADW onset determination was based on Li *et al.* (2010) [77]. It affirms that the onset of an ADW is related to a zero – crossing point before a maximal inflection of its derivative. A third order low pass Bessel filter is applied to the raw signal, which a cutoff frequency of 30 Hz to remove noises and artifacts that are common in unprocessed signals. Then, differentials are calculated and the local extreme that corresponds to the maximal inflection point of each pulse is determined by applying a magnitude threshold. Finally, and with the maximal inflections points determined, the first zero crossing before each maximum is calculated, matching the onset for each ADW pulse [24].

5.2.2.2 Baseline noise removal

The baseline noise can be caused by electrical signal fluctuations, slow motion of the neck attached PZ probe, motion artifacts or due to patient breathing. If this noise isn't properly removed, there will be an inaccurate determination of the characteristic points. Also, real time visualization results are improved without the baseline noise.

The PZ probe has real time baseline elimination based on a reliable baseline restorer [14]. Our software also has a baseline fit algorithm for baseline wander removal. This algorithm determines the baseline index points, and interpolates the baseline fit from the baseline indexes. The signal is corrected by vertical adjustment of each sample between two successive baseline indexes and matching the baseline point index to zero amplitude [24].

5.2.2.3 Morphological analysis

During acquisitions, motion artifacts originated from voluntary and involuntary subject movement causes volume changes in the ADW, and so, abnormal heartbeats can arise. After baseline removal, each ADW beat is analyzed, the mean ADW is calculated, and the RMSE between the mean ADW pulse and every individual pulse is calculated, and each pulse is computed. The criteria used to remove these pulses consisted in the presence of wide variations in amplitude and width pulse [24].

5.2.2.4 Pulse segmentation and normalization

After the morphological analysis and abnormal pulse flagging, we should have a clean ADW signal. This signal is segmented into individual pulses while the abnormal flagged pulses are removed. As the acquired ADW is not calibrated, only morphological characteristics of the ADW are analyzed. Normalization procedures are needed to scale heterogeneous beats, so that their morphology can be compared efficiently [24].

5.2.2.5 Spatial feature extraction

Diverse characteristic points can be extracted from each individual ADW pulse. The spatial feature extraction was based in the method proposed by Almeida et al. (2011) [25]. With this method, several ADW characteristic points were determined, enabling the possibility to classify the ADW type and to establish other parameters to be used in the future.

6. Methodology

In this chapter, the work oriented procedures for ADW acquisition, ADW processing and posterior data pre – processing are described. The collected data was included in a database and different datasets were created to ultimately reach different objectives: dataset I will be used for repeatability assessment of the non – invasive system, dataset II will focus on a case study, and dataset III and IV will be used for classification and clustering methods, respectively. Demographic data for all the datasets are presented as well.

6.1 ADW acquisition

One hundred and fifty five volunteers (95 female), aged between 18 and 80 years were included in this study. Subjects were mainly recruited from advertisements placed in public platforms calling for healthy volunteers and from *Centro Hospitalar e Universitário de Coimbra* for unhealthy volunteers. The research was approved by the Committees of the *Centro Hospitalar e Universitário de Coimbra.*

Before ADW acquisition, all subjects gave informal consent after full explanation of the purpose, nature and risk of the used procedures. Subject additional data was registered: age, gender, weight, height, smoking habits and diabetes history were collected, and body mass index (BMI) was calculated for all subjects. SBP, DBP and HR values were measured with an automatic digital oscillometric sphygmomanometer (Omron M6 Confort, Kyoto, Japan).

In a few cases (especially in unhealthy subjects), obtaining SBP, DBP and HR could not be possible. And so, the guidelines in Table 6.1 were used to correct missing data. The age and the subject's healthy or unhealthy status were considered in completing SBP, DBP and HR missing data. Age was divided in two groups for healthy subjects, < 40 and > 40, as SBP usually increases with aging in healthy subjects. For unhealthy subjects, it was acknowledged if they were in critical condition, or not. HR correction was always considered to 75 beats per minute (BPM), as it is the value for which AIx is usually corrected when assessing AIx @ 75 BPM (although AIx was not corrected in this work). References values for SBP and DBP were taken from [78].

The patient measurement protocol was mainly based on the subject condition standardization of Van Bortel *et al.* (2002) [79]. Subjects remained quiet and seated on a comfortable chair. Acquisitions were made at a similar time of the day, and in the same temperature controlled room (22 – 23ºC), to minimize the climatic variation. The non – invasive PZ sensor is placed over the carotid artery during data acquisition and held by a collar to clench the sensor in the patient's neck, and to rule out the influence of the operator, because interaction of the sensor with the operator's hand could increment additional noise. At least 3 acquisitions of roughly 30 – 40 seconds were made per subject.

If differences between left and right carotid in a subject was studied, 3 acquisitions of at least 30 seconds were performed for each carotid. If repeatability assessment between months is performed, the subject would return roughly 1 month later to perform the acquisition once again, undergoing the same procedures and conditions as in the first monthly acquisition.

Table 6.1 – Guidelines for SBP, DBP and HR missing data, considering healthy and unhealthy subjects;

Missing Variable	*Healthy (< 40 years)*	*Healthy (> 40 years)*	*Unhealthy*	*Unhealthy (critical condition)*
SBP (mmHg)	110	120	140	160
DBP (mmHg)	75	80	90	100
HR (bpm)	75			

6.1.1 ADW signal reproducibility

After ADW acquisitions, ADW processing routines were performed for each measurement. ADW signal reproducibility was assessed after pulse segmentation methods to illustrate the low variability of this non – invasive method. 4 subjects with different ADW type predominance were chosen. For each ADW type, two consecutive waveform and five consecutive waveforms were extracted, and the average waveform was plotted as tested for correlation.

6.2 Data pre - processing

Before we analyze the data, it must be rigorously pre – processed to avoid low – quality results. Usual data cleaning routines (missing value removal and outlier identification) were executed to discover and correct discrepancies in data. The gain ratio attribute evaluator technique for irrelevancy and redundancy analysis was performed before data mining procedures. Data normalization to transform the attributes into a 0 to 1 range were applied to the data in neural network classification tasks, due to their need of standardized inputs to perform efficiently.

In missing value removal routines, all type D ADWs (see section 2.3.2) were removed from the database. This was due a technical issue of the spatial feature extraction algorithm, which is not able to detect a reflection point (which is actually an intrinsic characteristic of the type D waveform) and, therefore, AIx can't be calculated. As the reflection point and AIx values of the type D ADWs are equal to zero, the non – removal of these waveforms would eventually present inaccurate results of the subsequent descriptive statistics and data mining routines. And so, only ADWs from type A, B and C were stored in the database.

6.3 Database

After all data pre – processing routines were performed, the final result is a group of pulse tuples with a maximum of 29 features for each subject in both *.mat* and *.arff* format, and all the processed data are stored in a dedicated database. From all the subjects, roughly 80% are healthy subjects between 18-30 years old, 5% are healthy subjects between 30 – 70 years old and 15% are unhealthy subjects. 8% of the volunteers have data in for left/right and two successive months, for repeatability assessment reasons.

We considered healthy volunteers as subjects with no documented history of CV disorders. Unhealthy subjects were classified according to their hypertensive status and CV disorder history. Age was not a discriminant factor in this manual classification.

6.3.1 Feature characterization

Each tuple is in fact, a previously parameterized pulse, and has a maximum of 29 attributes (plus 1 attribute in the *.arff* file, which represents the desired class). Prominent points extracted from the ADW were selected: systolic point time (SP_T), reflection point time (RP_T), dicrotic wave time (DW_T), systolic point amplitude (SP_A), reflection point amplitude (RP_A) and dicrotic wave amplitude (DW_A). AIx was calculated according to the guidelines in Table 3. Other ratios of interest (R_1 – R_4) were calculated using the feature points extracted values. Some statistical measurements were performed to assess the variance associated to pulse morphology. For time and amplitude positions, the root mean square of successive differences (RMSSD) for each of prominent points. The full width at half maximum (FWHM) was also calculated for all pulses.

Demographic data that was acquired after informal consent (age, gender, smoker, diabetes, height, weight, BMI, SBP, DBP, and HR) complete the feature composition. When applicable, the dataset class is classified as: 1 = healthy subjects, 2 = unhealthy subjects. Each attribute is further described and detailed in Table 6.2.

Table 6.2 – Description of all the attributes included in the database. Attributes with white background were used for dataset I and II. Attributes with white and brown background were used for dataset III. All the attributes on the table were used in dataset IV:

Attribute	*Description*	*Equation*	*Unit*
SP$_T$	Time at systolic time / upstroke time	--	[ms]
RP$_T$	Time at reflection point		
DW$_T$	Time at dicrotic wave		
SP$_A$	Systolic amplitude		[a.u]
RP$_A$	Reflected wave amplitude		
DW$_A$	Dicrotic wave amplitude		
AIx	Augmentation Index	*See chapter 2.4.3 (Table 2.2)*	[%]
R1	Downstroke time between systolic and dicrotic wave	$\|SP_T - DW_T\|$	[ms]
R2	Quotient between dicrotic wave amplitude and systolic amplitude	DW_A/SP_A	[a.u]
R3	Difference between systolic amplitude and reflected wave amplitude	$\|SP_A - RP_A\|$	[a.u]
R4	Quotient between systolic amplitude and reflected wave amplitude	$\begin{cases} RP_A/SP_A \ if \ SP_T < RP_T \\ - RP_A/SP_A \ if \ SP_T > RP_T \end{cases}$	[a.u]
FWHM	Full width at half maximum		[ms]
RMSE	Root mean square error between each pulse and the average pulse (pulse morphology variability)	$\sqrt{\frac{\sum_{i=1}^{n-1}(x_{1,i} - x_{2,i})^2}{n}}$	[%]
RMSSD_SP$_T$	Root mean square of successive differences of attribute X (RMSSD$_X$), with X = SP$_T$ $\vee$ RP$_T$ $\vee$ DW$_T$ $\vee$ SP$_A$ $\vee$ RP$_A$ $\vee$ DW$_A$	$\sqrt{\frac{\sum_{i=1}^{n-1}(x_{i+1} - x_i)^2}{n-1}}$	[ms]
RMSSD_RP$_T$			
RMSSD_DW$_T$			
RMSSD_SP$_A$			[a.u]
RMSSD_RP$_A$			
RMSSD_DW$_A$			
Age		--	[years]
Gender	1 = Male 2 = Female	--	[male/female]
Smoker	1 = Smoker 2 = Non - Smoker	--	[yes/no]
Diabetes	1 = No 2 = Yes	--	[yes/no]
Body height		--	[m]
Body weight		--	[kg]
BMI	Body mass index	$\frac{Body\ weight}{(Body\ height)^2}$	[kg/m^2]
SBP	Systolic blood pressure	--	[mmHg]
DBP	Diastolic blood pressure	--	[mmHg]
HR	Heart rate	--	[beats per minute]

The work – oriented methodology from section 6.1 to 6.3 is expressed on Figure 6.1 in a concise graphical manner.

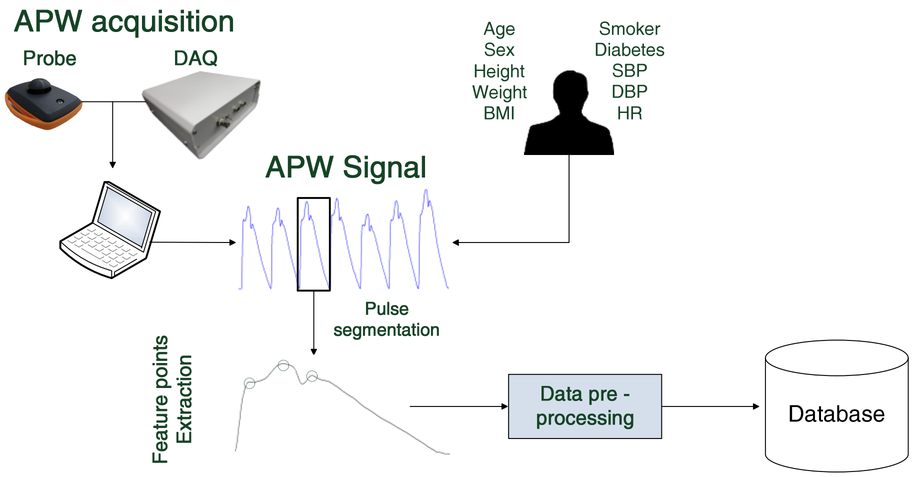


Figure 6.1 – Work – oriented methodology;

6.4 Experimental datasets

In the database, each subject tuples can be posteriorly merged with other subject's tuples automatically with the help of a tool, and without overlapping. This is crucial in creating different datasets without a time – consuming manual merging and for studying different aspects of scientific interest. Therefore, four different experimental datasets were created: dataset I was created for repeatability assessment and dataset II focused on a case study of a group of subjects which were monitored under carotid intervention. Dataset III and IV were created for data classification and data clustering procedures, respectively. Demographic data comparison is expressed in Table 6.3.

Table 6.3 – Demographic data comparison for each of the created datasets. Data are expressed as mean ± SD;

Variable	*Dataset I*	*Dataset II*		*Dataset III*		*Dataset IV*
			Sub – group I	*Sub – group II*	*Diagnostic*	
Age (years)	23.50 ± 2,43	72.5 ± 5.44	24.16 ± 3.86	58.16 ± 1..77	35,16 ± 11,77	21.19 ± 2,28
Gender (M/F)	4/8	4/2	13/12	12/13	4/6	31/62
Smoker (yes/no)	0/13	M.D*	3/22	4/21	1/9	7/86
Diabetes (yes/no)	0/13	M.D*	0/25	2/23	M.D*	0/93
Weight (kg)	59.66 ± 10.96	M.D*	65,28 ± 10,42	72.66 ± 14.13	M.D*	61.53 ± 10.25
Height (m)	1.66 ± 0.06	M.D*	1,70 ± 0,06	1.64 ± 0.09	M.D*	1.68 ± 0.09
BMI (kg/m^2)	21.45 ± 2.71	M.D*	21,76 ± 4,89	29.08 ± 8.64	M.D*	21.62 ± 2.63
SBP (mmHg)	105.25 ± 5.86	144.33 ± 40.08	110,20 ± 11,94	150.84 ± 26.29	M.D*	108.26 ± 11.88
DBP (mmHg)	65.66 ± 6.71	81 ± 18.34	69,80 ± 10,17	88.92 ± 16.45	M.D*	69.54 ± 7.64
HR (beats/min)	71.16 ± 10.64	72.5 ± 3.53	68,44 ± 10,72	63.04 ± 7.71	M.D*	70.84 ± 10.87

* M.D – missing data

6.4.1 Dataset I – Repeatability Assessment

Dataset I is composed of 12 healthy subjects (8 female), aged < 30 years and with no documented history of CV disorders. Two successive monthly sessions containing both left and right carotid data were monitored. As referred in 6.1, at least 3 acquisitions of roughly 30 – 40 were made per carotid on each subject, for a total of 3 trials per carotid site, per month. In this group, only 7 attributes were used: SP_T, RP_T, DW_T, SP_A, RP_A, DW_A and AIx. This dataset has a total of 2983 pulses, with ≈ 200 – 250 pulses for each subject.

In this dataset, the differences between trials (intra – subject variability), left and right carotid and between successive months for each subject were studied using statistical approaches, with the objective of assessing the repeatability of the developed non – invasive system in acquiring data.

All data were analyzed with Predictive Analytics Software Statistics 18.0 (SPSS, Inc, Chicago, IL). The level of statistical significance was set at $p < .05$ for all analyses. A Bland - Altman GUI was developed to create, visualize and save Bland – Altman plots.

6.4.1.1 Normality assessment

The first step when dealing with new data is the normality analysis, to verify if there is any violation of the statistical techniques assumptions, which will ultimately decide in using a parametric or non – parametric test in advanced statistical analysis. All attributes were tested for normality using the KS, and divided by month, trial, carotid site and subject. If the significance result of the KS test is < .05, the null hypothesis is rejected, and the variable has a non – normal distribution. Else, the variable is normal.

6.4.1.2 Correlation analysis

Correlation analysis was performed to assess if linear relationships do exist among the seven variables. Scatterplots were constructed to primarily visualize possible relationships and to identify the existence of sub – groups in the dataset. With respect of the given result of our distributions (6.4.1.1), parametric and/or non – parametric correlation tests will be used to assess the strength and direction of possible correlation.

6.4.1.3 General statistical analysis

After the normality tests, and with the assistance of quantitative measures and boxplot graphics, a global view from our data divided by month and carotid site is checked. A categorization of the data by subject is also important, as pulse variability may be different from subject to subject. Sub – categorization by month, carotid site and trial are done for each subject. The objective of this general statistical analysis is to assess differences between groups.

Statistical tests were used to assess significant differences in AIx, for each subject, according to different sub – categorizations: between months, between left and right carotid and between trials. A low rate of significant differences between groups can be a landmark in validating the repeatability of the non – invasive system. With respect of the given result of the normality assessment, parametric and/or non – parametric approaches were used to assess the differences between groups.

6.4.1.4 Agreement - Bland – Altman plots

Bland – Altman plots were constructed only for the AIx attribute, as it is the hemodynamic index of interest. The objective is to assess the agreement between two measurements with the non – invasive system, which will ultimately show the repeatability of the AIx measured by the PZ probe. Each Bland – Altman plot will have 12 points, as each point represents a subject. The following agreements tests were conducted:

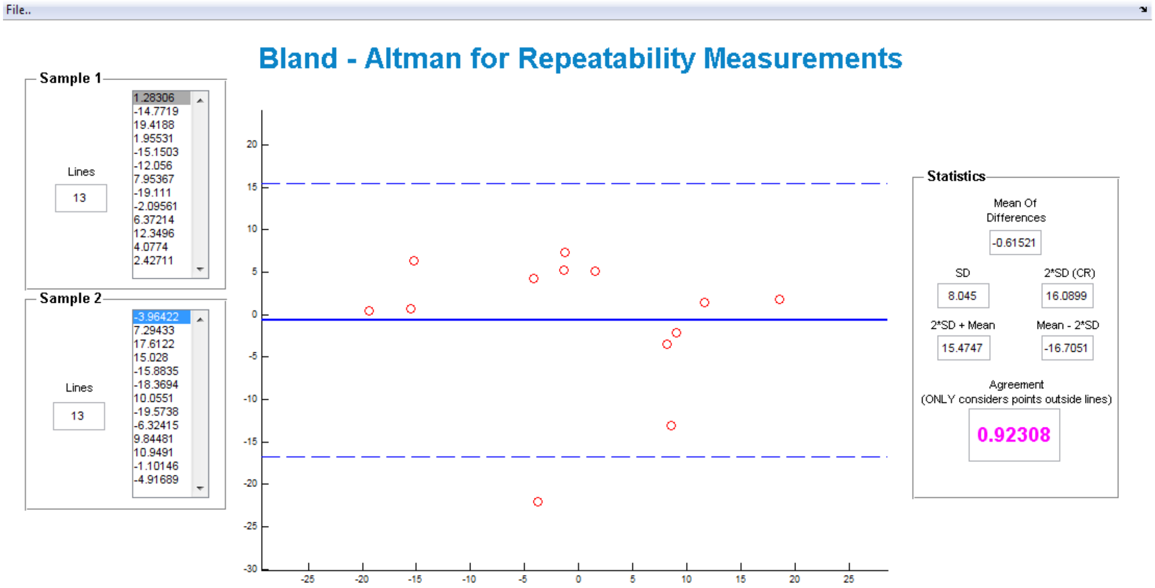
- Test A - Month 1 & month 2
- Test B - Month 1 right carotid & month 1 left carotid
- Test C - Month 2 right carotid & month 2 left carotid

6.4.1.4.1 <u>Bland – Altman developed GUI</u>

A Bland – Altman GUI was developed in MATLAB during the course of this work due to difficulties in finding an efficient compatible Bland – Altman software. The software can measure the agreement between two variables with the same number of lines and one single column. The program only supports .mat – type files. If more than one column exists on the variable, the program considers the first column. If both samples have a different number of lines, the program stops.

Figure 6.2 shows the software with the Bland – Altman graphic displayed at the center and with more information displayed in the left and right part of the GUI. The left part (Sample 1 and Sample 2) displays each value of each variable, for direct quantitative comparison. The right part (Statistics) gives important quantitative information about the plotted Bland – Altman graphic, including the mean of differences, the SD, the coefficient of repeatability (CR, which is the same as 2*SD, the 2*SD + Mean and Mean – 2*SD limits and finally, the agreement, which is expressed as ratio, between 0 and 1. We expect 95% to be inside the limits, and so, an agreement value higher or equal to 0.95 will reveal good agreement. If somehow, any point is outside the limits (very common), the agreement will lower progressively.

The software also supports saving to .bmp image, or to .fig (MATLAB$^{®}$ Figure) and statistic values are saved in a *.xls* file (Microsoft Excel$^{®}$ compatible).



Bland - Altman for Repeatability Measurements

Sample 1

Lines: 13

1.28306
-14.7719
19.4188
1.95531
-15.1503
-12.056
7.95367
-19.111
-2.09561
6.37214
12.3496
4.0774
2.42711

Sample 2

Lines: 13

-3.96422
7.29433
17.6122
15.028
-15.8835
-18.3694
10.0551
-19.5738
-6.32415
9.84481
10.9491
-1.10146
-4.91689

Statistics

Mean Of Differences: -0.61521

SD: 8.045

2*SD (CR): 16.0899

2*SD + Mean: 15.4747

Mean - 2*SD: -16.7051

Agreement (ONLY considers points outside lines): 0.92308

Figure 6.2 – Bland – Altman for Repeatability Measurements software, with the Bland – Altman plot in the display;

6.4.2 Dataset II – Case Study: Angioplasty

In this group, the same attributes as dataset I were used: SP_T, RP_T, DW_T, SP_A, RP_A, DW_A and AIx. Smoker, diabetes, weight, height and BMI information was not collected, as it was not necessary for the objective of this dataset.

Subjects from dataset II suffered from stenosis and were monitored under angioplasty carotid intervention, in collaboration with *Centro Hospitalar e Universitário de Coimbra*. An angiography was carried out to visualize the blood circulation. The ADW was simultaneously collected with the developed PZ probe and the cardiac catheterization monitor *Axiom Sensis* (Siemens, Munich, Germany) invasive equipment. A small segment of 3 – 4 seconds was chosen from each method before and after carotid intervention. A direct comparison was performed, although it was only possible through visual inspection, due to limitations in the raw data availability that was obtained from the invasive device.

6.4.3 Dataset III – Data Mining: Classification

Dataset III has three sub – groups included. The first sub – group is composed of 25 healthy subjects, while the second sub – group is composed of 25 unhealthy subjects. Both of these sub – groups were used for classifier construction.

All attributes were used, except demographic data, for a total of 19 attributes (which class label totalizing 20). Categorical class values were determined for each sub – group, with the first sub – group having class = 1 (healthy) and the second sub – group having class = 2 (unhealthy). Considering only the first and second sub – group, dataset III contains a total of

2947 pulses. Attributes were normalized previously to any classification procedure, to speed up the learning phase and rule out weight problems (which are common in ANN).

The third and last sub – group consists of undiagnosed subjects, in other words, subjects that weren't used for classifier construction and have unknown class. 10 subjects that weren't included in classifier construction compose this diagnostic sub – group and this multiple classification methodology will be tested against this undiagnosed sub – group to predict their CV condition.

As we have pre – determined classes in dataset III, it is an ideal dataset for classification procedures, in other words, constructing models that describe and distinguish between healthy and unhealthy subjects, for the purpose of being able to use the classifier to predict the class of subjects whose class label is unknown. For these classification procedures, Weka 3.6.4 (Waikato Environment for Knowledge Analysis), which is a free machine learning software based in Java^{TM} language, was the preferred option, due to its affordability, versatility and efficiency. Weka supports *.arff* files (attribute - relation file format), which is one of the output files of the ADW Pulse Analzyer after ADW processing.

6.4.3.1 Classifier selection

Eight classifiers, from four different types of classification methods, are selected to be constructed. Classifiers are evaluated according to their accuracy, sensitivity, specificity, precision and training speed. RMSE is also calculated for each classifier. The classifiers are:

- **Decision tree induction -** C4.5 (which is named J48 in Weka) and random forest;
- **Associative classification -** RIPPER (named JRip in Weka);
- **Bayesian classification -** Naïve – Bayes and Bayesian network;
- **Neural networks** – 1 – hidden layer MLP, 2 – hidden layers MLP and RBF;

The three classifiers with best combined result of accuracy, specificity and sensitivity are chosen for the following multiple diagnostic procedure.

6.4.3.1.1 <u>ANN performance study</u>

Special focus was given to neural network methods, as they were never tested before among the previously performed classification procedures with the developed non – invasive system. Also, neural networks are quite sensitive to slight parameter changes, and so, a specific parameter tuning was performed for each neural network – based classifier, ensuring a maximization of the classifier's accuracy in differentiating healthy from unhealthy subjects. Different parameter configurations were tested for each neural network, while manipulating one of the parameters and comparing the accuracy, sensitivity, specificity and precision measures for each configuration. The best configurations are chosen for classifier comparison with decision tree induction, Bayesian and associative classification methods.

During this neural network parameter tuning, all nominal attributes were removed, as neural networks cannot handle nominal values as inputs, but only as targets. In neural networks, the dataset was also normalized between 0 and 1.

6.4.3.2 Diagnostic with multiple classifier methodology

A multiple classifier methodology is used due to its potentialities in producing more reliable results in comparison with single classifier analysis. These ensemble methodologies consist on the assumption that the classification should not be based on the result of one, and only one classifier. This class prediction should be done by a combination of more than one classifier, minimizing the possibility of occurring FPs and FNs.

A weight voting classification is adopted, and the 3 - best classifiers are assembled to predict the class of subjects. Considering that the voting system is equal – weighted for the three classifiers, if at least two classifiers consider the subject as healthy, the predicted class is healthy. Else, high CV risk factor/unhealthy is the given classification.

6.4.4 Dataset IV – Data Mining: Clustering

This dataset consists of 93 healthy and young subjects between 18 – 30 years, with no documented history of CV disorders with unknown class label. A total of 4471 pulses compose this dataset. All of these subjects' data were obtained at two different moments: in March, while collaborating with *SCDSOS* (Sudden Cardiac Death Screening of Risk Factors) and in May, during a mobilized screening in the Department of Physics of the University of Coimbra in collaboration with MsC student Inês Santos. This dataset is destined for clustering methods, which will attempt to categorize the objects of the dataset into two and three different clusters, where each cluster will represent different risk groups based on arterial patterns. The full list of 29 parameters was initially used for clustering procedures (with the unknown class attribute totalizing 30).

The objective of the clustering method in the context of this dissertation is to find and distinguish different risk groups in terms of future CVD development in a healthy dataset.

Just like in classification routines, Weka was the preferred software for clustering. Only two clustering methods were used: k – means algorithm and EM, due to Weka's software limitations.

6.4.4.1 Two and three risk groups clustering

EM and k – means clustering were used for partitioning the dataset in two distinct risk groups. Results were compared for each clustering method. For three risk groups partitioning, only the algorithm with more satisfying results was used.

When k – means clustering is performed, the categorical attributes (gender, smoker, diabetes) were removed, due to k – means incompatibility with these features. So, for EM, 29 attributes were used, and 26 attributes were used in k – means clustering.

6.4.4.1.1 Attribute subset selection

After the group clustering, attribute subset selection with the two partitioned class was performed to find the most important attributes in class determination for two and three risk groups. The gain ratio attribute evaluator with 10 – fold CVN attribute selection mode is the preferred attribute subset selection method.

7. Results & Discussion

All the results for each dataset are presented this chapter. This section also presents inter – communication improvements between ADW acquisition and ADW processing GUIs and ADW signal reproducibility during acquisition.

An extensive discussion is presented in the last sub – section of each dataset section.

7.1 ADW acquisition

ADW acquisition procedures went underway according to the guidelines presented in 6.1. A minor but important modification is now presented, due to its importance in ADW acquisition routines. Signal reproducibility during acquisition is also assessed.

7.1.1 Inter – communication between GUIs

Even though the previously developed GUIs for ADW acquisition and ADW processing are fully functional, there was no direct communication between them. This evoked a big problem: the impossibility to confirm in the moment if the acquired signal was good for subsequent processing and feature points extraction, as sometimes, due to intrinsic or external factors, it was difficult to assess if a collected signal was well acquired by the platform, or not. This problem was solved by adding and programming a button that is able to transfer all the raw information from one GUI to another.

In the Figure 7.1, it is possible to visualize the ADW Acquisition Module and the ADW Pulse Analyzer Module. There's a highlight in the "Save&Transfer" button of the ADW Acquisition Module, which is responsible for transferring all the current acquired data to the ADW Pulse Analyzer Module. In the ADW Pulse Analyzer Module, the "Transfer from RS232" button is highlighted, and will successfully transfer the recently acquired raw signal for ADW analysis.

With this minor adjustment, performing quality control procedures right after the ADW acquisition is directly feasible and, therefore, there is an improvement in the quality of parameter data stored in the database.

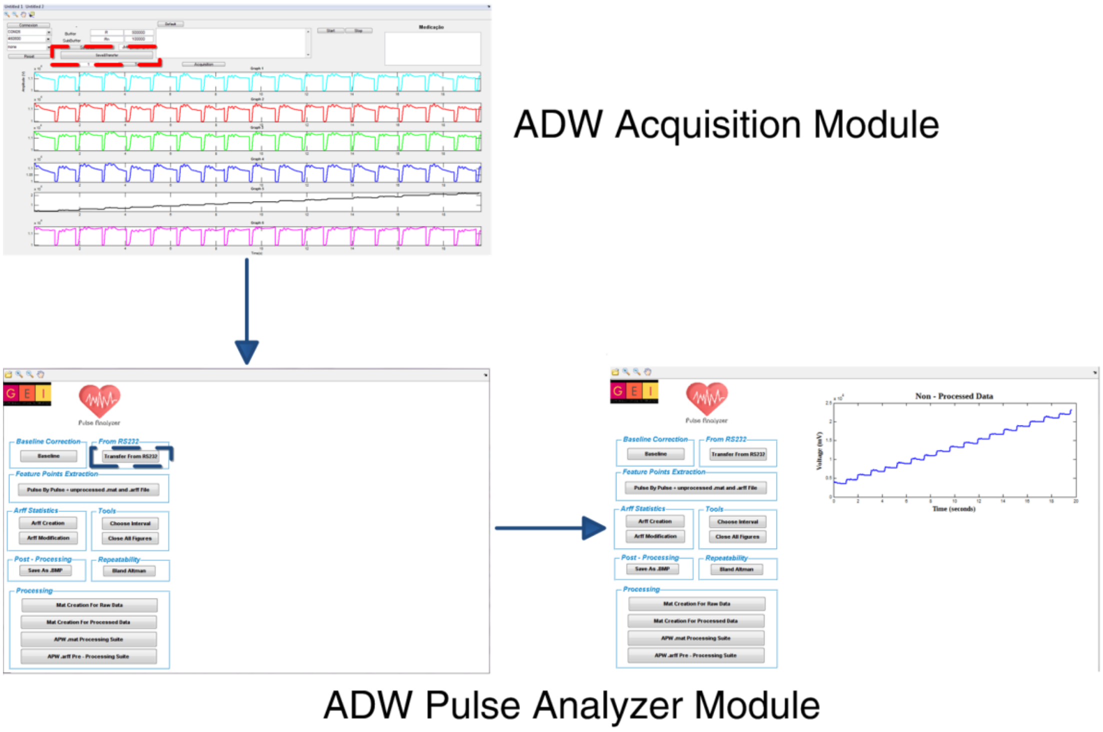


Figure 7.1 – Communication between the APW Acquisition and APW Pulse Analyzer modules;

7.1.2 ADW signal reproducibility

The next image shows two consecutive normalized ADWs from the same type, and from the same subject (Figure 7.2). Correlation results are also expressed in the figure. The subject selection in the database was random for each ADW type.

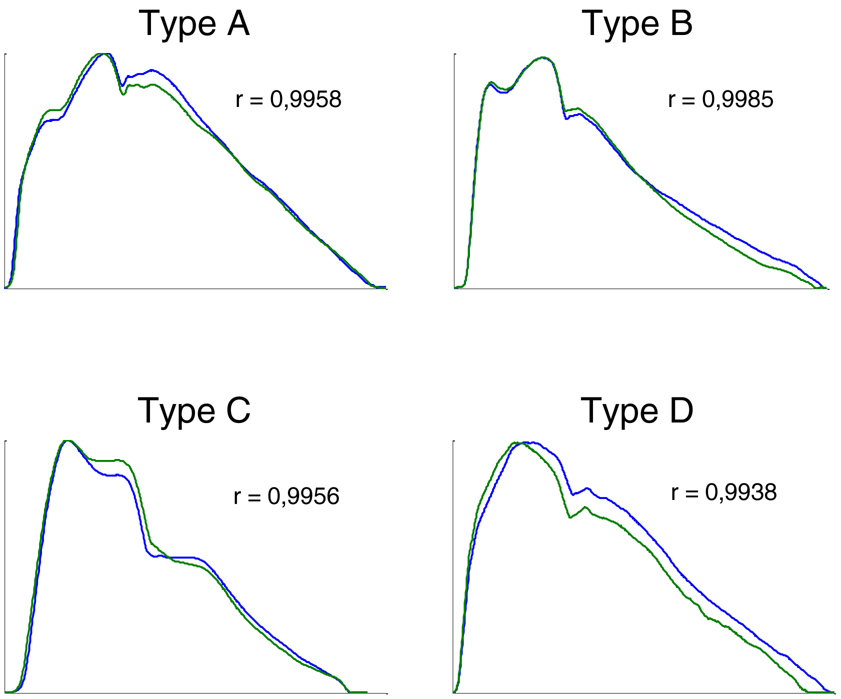


Figure 7.2 – Signal reproducibility for each ADW type;

Each ADW type has shown excellent reproducibility, with good correlation ($r > 0.99$) for all ADW types. Coefficient of determination (r^2) values for each case are > 0.985 (lowest r^2 corresponds to ADW Type D = 0.9876).

For each of the ADW types, 10 sets of two consecutive normalized pulses were randomly extracted for different subjects, and correlation coefficients were computed. Results are expressed in Table 7.1.

Table 7.1 – Correlation coefficient (r) results for 10 sets of two consecutive normalized pulses for each ADW type;

Set of 2 consecutive pulses	*ADW Type*			
	A	*B*	*C*	*D*
1	0.9932	0.9940	0.9945	0.9934
2	0.9935	0.9932	0.9923	0.9834
3	0.9942	0.9954	0.9965	0.9912
4	0.9897	0.9965	0.9956	0.9899
5	0.9940	0.9954	0.9923	0.9876
6	0.9953	0.9959	0.9920	0.9945
7	0.9912	0.9972	0.9935	0.9865
8	0.9916	0.9934	0.9956	0.9923
9	0.9935	0.9979	0.9941	0.9912
10	0.9932	0.9929	0.9922	0.9945

Globally, ADW type D has the lowest correlation coefficients, even though they are very satisfying. The lowest r^2 in the table is Run 2 of ADW type D ($r = 0.9834$), with $r^2 = 0.9670$.

Therefore, **it is safe to say that spatial feature information that is extracted from the ADWs will present low variability between pulses, for the same subject.**

7.2 Dataset I – Repeatability Assessment

To successfully validate the repeatability of our sensor, its variability needs to be evaluated according to each of the groups included in the study: month, carotid site and subject.

7.2.1 Normality assessment

The results of the KS test for assessing the normality of our data in terms of month, carotid site and trial are expressed in Table 7.2. Normality was also assessed for each subject, and results are represented in Table 7.3.

Table 7.2 – KS test significance values for each of the variables, categorized by month, carotid site and trial. Values in bold correspond to normal distributions. A P – value of < .05 was considered as significant;

Variable	*Total*	*Month 1*	*Month 2*	*Left*	*Right*	*Trial 1*	*Trial 2*	*Trial 3*
SP_T (ms)	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05
RP_T (ms)	< .05	< .05	< .05	< .05	< .05	< .05	**0.10**	< .05
DW_T (ms)	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05
SP_A (a.u)	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05
RP_A (a.u)	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05
DW_A (a.u)	< .05	< .05	< .05	< .05	< .05	< .05	**0.07**	< .05
AIx (%)	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05

The KS test for each of the categories (month, carotid site and trial) revealed that the dataset is non - normal in its majority. A normal distribution was presented in two occasional situations (and are represented in bold, on each table). When the KS test was performed with subject as a category (Table 7.3), results still indicate a non – normal distribution for every attribute, in overall. However, DW_A and RP_T evidenced normal distribution in some subjects.

Concluding the normality assessment of dataset I, **the data presented here is almost 100% non – normal.** This result was expected, as our dataset I is relatively large, and non – normal distributions are quite common in large datasets. These results might suggest that a latter non – parametric approach when using statistical tests to assess differences between groups might be a wiser choice, because in non – parametric tests, the values are converted to ranks, so the actual distribution of the values in a variable does not matter.

Table 7.3 – KS test significance values for each of the variables, categorized by subject. Values in bold correspond to normal distributions. A P – value of < .05 was considered as significant;

Variable	*Subject*											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
SP_T (ms)	< .05	< .05	**0.17**	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05
RP_T (ms)	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05	**0.12**	**0.21**	**0.47**	< .05
DW_T (ms)	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05	**0.06**
SP_A (a.u)	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05	< .05
RP_A (a.u)	< .05	< .05	< .05	< .05	**0.10**	< .05	< .05	< .05	**0.26**	< .05	< .05	< .05
DW_A (a.u)	< .05	< .05	< .05	**0.13**	**0.19**	**0.21**	< .05	**0.10**	< .05	**0.45**	**0.25**	**0.14**
AIx (%)	< .05	< .05	< .05	< .05	**0.28**	< .05	< .05	**0.54**	< .05	< .05	< .05	< .05

7.2.2 Correlation analysis

Initially, and before performing any correlation tests, two scatter plots between some of the variables included in dataset I were constructed, for a primary correlation analysis. The scatterplots of RP_T as a function of SP_T and RP_T as a function of AIx are presented in Figure 7.3.

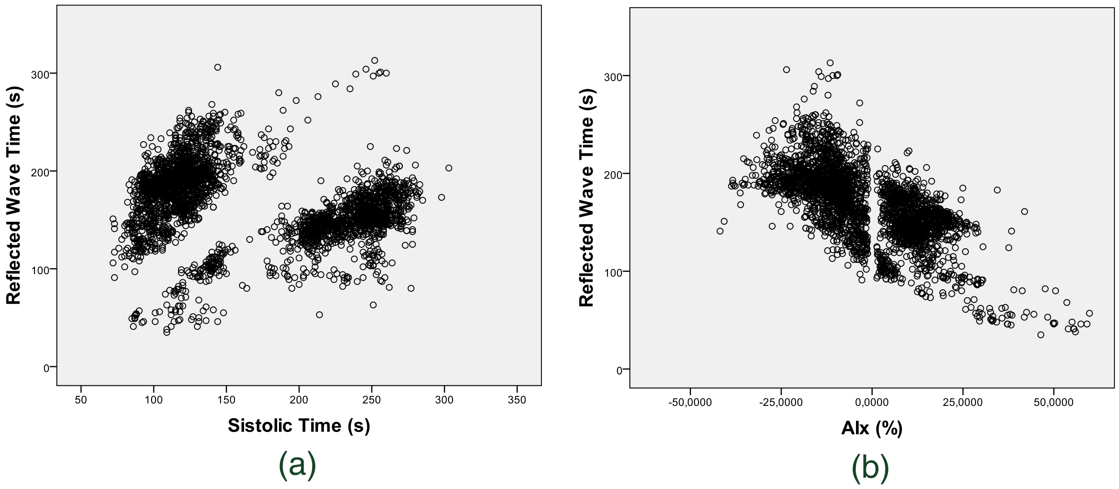


Figure 7.3 – Scatter plot of (a) RP_T as a function of SP_T; and (b) RP_T as a function of AIx;

In both scatter plots, it is possible to visualize the existence of two completely distinct groups. One of the groups represents pulses with negative AIx, which is related with late reflection wave arrival, and the other group represents pulses with positive AIx, directly related with early reflection wave arrival (and consequently, late systolic time point). **For that reason, the correlation analysis will be performed for two separate groups. In group A, all pulses will have negative AIx (which means that $RP_T > SP_T$). Group B will have the remaining pulses, which have positive AIx (where $RP_T < SP_T$).**

Correlation analysis with the Pearson's test was performed for both groups individually, to assess the strength and direction of possible linear associations between variables. Although the data is non – normal, dataset I is large enough to use a parametric alternative. Results of the correlation analysis with the Pearson's test for both groups are presented in Table 7.4. Guidelines by Cohen (1988) [59] will be used to determine if the correlations reveal strong, medium or low associative significance (see Table 3.2).

Table 7.4 – Analysis with the Pearson's product – moment correlation test for Group A and Group B. r values are expressed in each cell. Numbers on bold indicate strong correlation;

Group A Pearson's	***SP_T (ms)***	***RP_T (ms)***	***DW_T (ms)***	***SP_A (a.u)***	***RP_A (a.u)***	***DW_A (a.u)***	***AIx (%)***
SP_T (ms)		**0.618**	0.351	**0.557**	0.292	0.351	0.292
RP_T (ms)	**0.618**		0.405	0.370	-0.361	-0.131	0.361
DW_T (ms)	0.309	0.405		0.284	-0.027	-0.343	-0.027
SP_A (a.u)	**0.557**	0.370	0.284		0.412	0.336	0.412
RP_A (a.u)	0.292	-0.361	-0.027	0.412		**0.657**	**1.000**
DW_A (a.u)	0.351	-0.131	-0.343	0.336	**0.657**		0.657
AIx (%)	0.292	0.361	-0.027	0.412	**1.000**	0.657	

Group B Pearson's	***SP_T (ms)***	***RP_T (ms)***	***DW_T (ms)***	***SP_A (a.u)***	***RP_A (a.u)***	***DW_A (a.u)***	***AIx (%)***
SP_T (ms)		**0.775**	**0.927**	-0.005	0.091	-0.230	-0.091
RP_T (ms)	**0.775**		**0.750**	-0.131	0.487	-0.246	0.487
DW_T (ms)	**0.927**	**0.750**		-0.078	0.133	-0.326	0.133
SP_A (a.u)	-0.005	-0.131	-0.078		0.017	0.037	0.017
RP_A (a.u)	0.091	0.487	0.133	0.017		0.189	**-1.000**
DW_A (a.u)	-0.230	-0.246	-0.326	0.037	-0.189		0.189
AIx (%)	-0.091	0-287	-0.133	-0.017	**-1.000**	0.189	

Analysis will be given for each group in separate:

- **Group A –** This group represents all pulses with negative AIx. A strong linear correlation between SP_T and RP_T was determined. Other strong associations are found between RP_A and DW_A, and SP_T and SP_A. However, one of the associations revealed perfect positive association: RP_A and AIx. As RP_A values get smaller, AIx values will also be negatively smaller, even if it usually represents a good result in terms of arterial stiffness.
- **Group B –** This group represents all pulses with positive AIx. As in group A, a strong linear association between SP_T and RP_T was determined. A strong linear association between DW_T and RP_A can also be visualized. In group B, instead of a perfect positive correlation between RP_A and AIx, the result is a perfect negative correlation. In other words, when RP_A is lower, AIx is subsequently higher (which indicates high arterial stiffness).

This perfect association between RP_A and AIx in both groups separately is not visible when both groups are coupled. This is logical, as RP_A only indicates the strength of the AIx. The direction is determined by the comparison between SP_T and RP_T, with a negative AIx occurring when $RP_T > SP_T$ and a positive AIx arising when $SP_T > RP_T$. If one would try to assess the correlation between RP_A and AIx with groups coupled, a very low association would be presented, as the direction of the RP_A values was not preserved when AIx values rose.

7.2.3 General statistical analysis

A comprehensive quantitative and visual statistical analysis follows, complemented with the use of statistical tests, if needed to confirm a hypothesis.

7.2.3.1 Categorization by month

While giving a quick overview of dataset I, Table 7.5 also compares dataset's I descriptive data categorized by month. A Mann – Whitney U non – parametric test was conducted to compare each of the continuous variables between months. Also, and as we are dealing with non – normal data, an AIx boxplot was constructed to visualize the median and the range in a simple and visual manner (Figure 7.4).

At a first look, it is possible to verify the existence of big differences between SP_T between month 1 and month 2, with month 1 having lower SP_T. RP_T only changed slightly, and so we will have a lower AIx value in month 1 than in month 2. As for SP_A, RP_A and DW_A, they seem to reveal very low variability, mainly because they are normalized attributes**. However,**

the Mann – Whitney test revealed that there are significant differences between months, in all seven variables.

Table 7.5 – Descriptive data categorized by month. All data is expressed as mean ± SD. P indicates the Mann – Whitney test significance value. A P – value of < .05 was considered as significant;

Variable	*Total*	*Month 1*	*Month 2*	*P*
SP_T (ms)	166.27 ± 62.44	155.63 ± 58.76	177.74 ± 64.26	< .05
RP_T (ms)	167.70 ± 37.36	168.90 ± 40.12	166.41 ± 34.09	< .05
DW_T (ms)	289.47 ± 52.27	285.98 ± 61.72	293.23 ± 39.31	< .05
SP_A (a.u)	0.996 ± 0.008	0.996 ± 0.008	0.996 ± 0.009	< .05
RP_A (a.u)	0.87 ± 0.08	0.86 ± 0.085	0.88 ± 0.066	< .05
DW_A (a.u)	0.71 ± 0.14	0.70 ± 0.15	0.72 ± 0.13	< .05
AIx (%)	-1.16 ± 15.05	-2.67 ± 15.94	0.48 ± 13.84	< .05

The boxplot of Figure 7.4 also displays a lower AIx for the first month. The median AIx for the first and second month is -6.66% and 3.89%, respectively, which represents a huge difference, in overall.

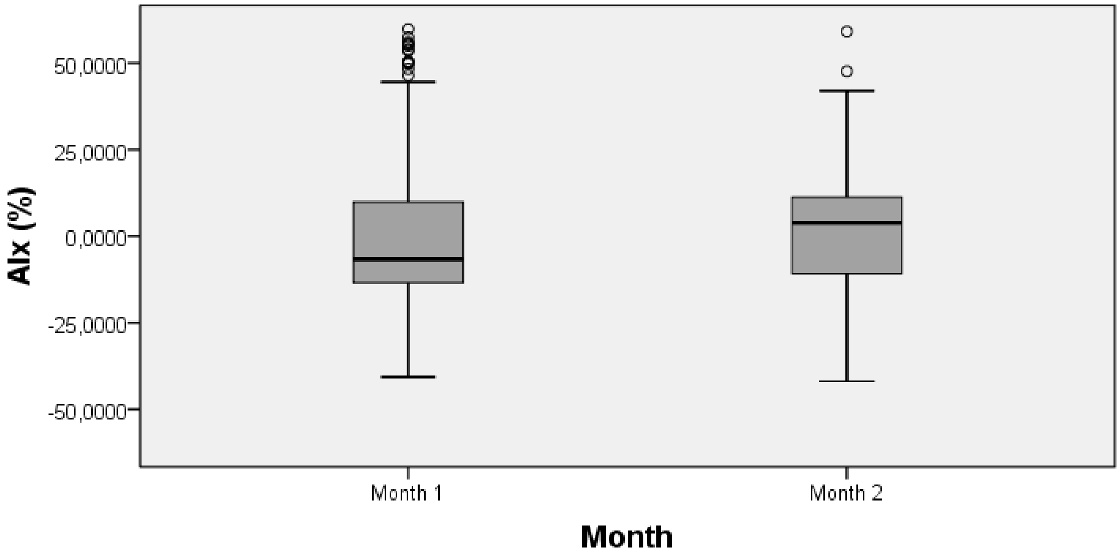


Figure 7.4 – Boxplot of the AIx categorized by month. The median for the first measurement is -6.66% and 3.89 for the second measurement;

Outliers can be having a huge impact in these dissimilarities between months, but as was referred before, in the methodology, no outlier elimination was executed, as it was hypothesized that they could have special meaning. It should be also reminded that this table only gives a global overview of the variability of our data between months, and subjects need to be analyzed individually. Even if the dataset's difference between months is significant, that does not mean that each subject has significant differences between months. A categorization by subject, sub – categorized by month will be later performed.

SP_A was removed from subsequent statistical analysis, as it was previously normalized to values ≈ 1, during pre – processing. Hence, its relevance in the analysis is negligible.

7.2.3.2 Categorization by carotid site and month

Table 7.6 and Figure 7.5 AIx boxplot display the differences between left and right carotid in each month. A Mann – Whitney non – parametric test was performed to compare each of the variables carotid site differences in each month.

Table 7.6 – Descriptive data categorized by month and sub – categorized by carotid site, to assess differences between left and right carotid in each month. All data is expressed as mean ± SD. P columns indicate the Mann – Whitney test significance value. A P – value of < .05 was considered as significant;

Variable	*Total*	*Month 1*			*Month 2*		
		Left	*Right*	*P*	*Left*	*Right*	*P*
SP_T (ms)	166.27 ± 62.44	159.02 ± 63.44	153.14 ± 55.15	**0.85**	178.52 ± 64.28	176.94 ± 64.27	**0.49**
RP_T (ms)	167.70 ± 37.36	165.97 ± 36.867	171.04 ± 42.24	**0.11**	165.98 ± 32.56	166.84 ± 35.61	**0.16**
DW_T (ms)	289.47 ± 52.27	290.03 ± 65.34	283.01 ± 58.80	< .05	294.53 ± 35.78	291.91 ± 42.60	**0.15**
RP_A (a.u)	0.87 ± 0.08	0.85 ± 0.09	0.87 ± 0.08	< .05	0.88 ± 0.06	0.88 ± 0,07	**0.33**
DW_A (a.u)	0.71 ± 0.14	0.71 ± 0.15	0.70 ± 0.15	**0.10**	0.73 ± 0.12	0.73 ± 0,13	**0.81**
AIx (%)	-1.16 ± 15.05	-2.12 ± 17.44	-3.07 ± 14.75	**0.31**	0.58 ± 14.28	0.36 ± 13,38	**0.64**

A primary analysis, complemented with the Mann – Whitney statistical test reveals that there are no significant differences between left and right carotid in each month and, therefore, measuring on the left or right carotid may be irrelevant considering our dataset I population (< 30 years).

Figure 7.5 reveals almost identical median for each of the carotid site while sub – categorized in each month.

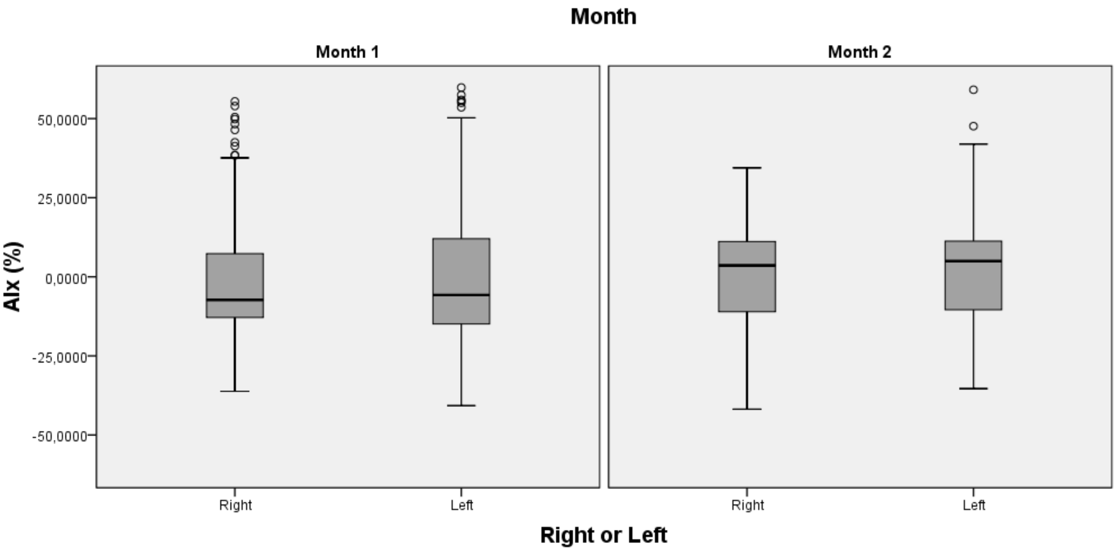


Figure 7.5 – Boxplot of the AIx categorized by carotid site and by month. For the first month, the AIx right and left carotid median is -7.31% and -5.73%, respectively. For the second month, the AIx right and left carotid median values are 3.58% and 4.96%, respectively;

When considering the first month, the AIx right and left carotid median is -7.31% and -5.73%, respectively. As for the second month, the AIx right and left carotid median values are 3.58% and 4.96%, respectively.

Finally, a Mann - Whitney U test was performed to find out if there are AIx differences between left and right carotid site, without month categorization. A P – value < .05 was considered as significant. Results demonstrated that there were no significant differences between left and right carotid site (U = - 1, 89, P = 0.06). Accordingly, the null hypothesis is not rejected. Although, this significance value is quite small and there is a higher risk of erroneously accepting the null hypothesis. The small significance value can be explained by the fact that both months were considered, and as assessed in sub – chapter 7.2.2.1, there are significant differences between months. An individual approach must be performed to assess this.

Concluding, no dissimilarities are found not only between measurements in right and left carotid in each month, but also between left and right carotid without month categorization.

7.2.3.3 Categorization by subject

An analysis of our data while categorized by subject seems like the most efficient way of describing the information at our disposal, as pulse variability may differ from subject to subject regarding each of the seven attributes and a particular analysis must be executed.

As AIx is the hemodynamic parameter of interest, it will be the only studied parameter in this section. Figure 7.6 displays an AIx boxplot, categorized by each of the twelve subjects that compose this dataset, and complemented with the IQR of the AIx of each subject (Table 7.7).

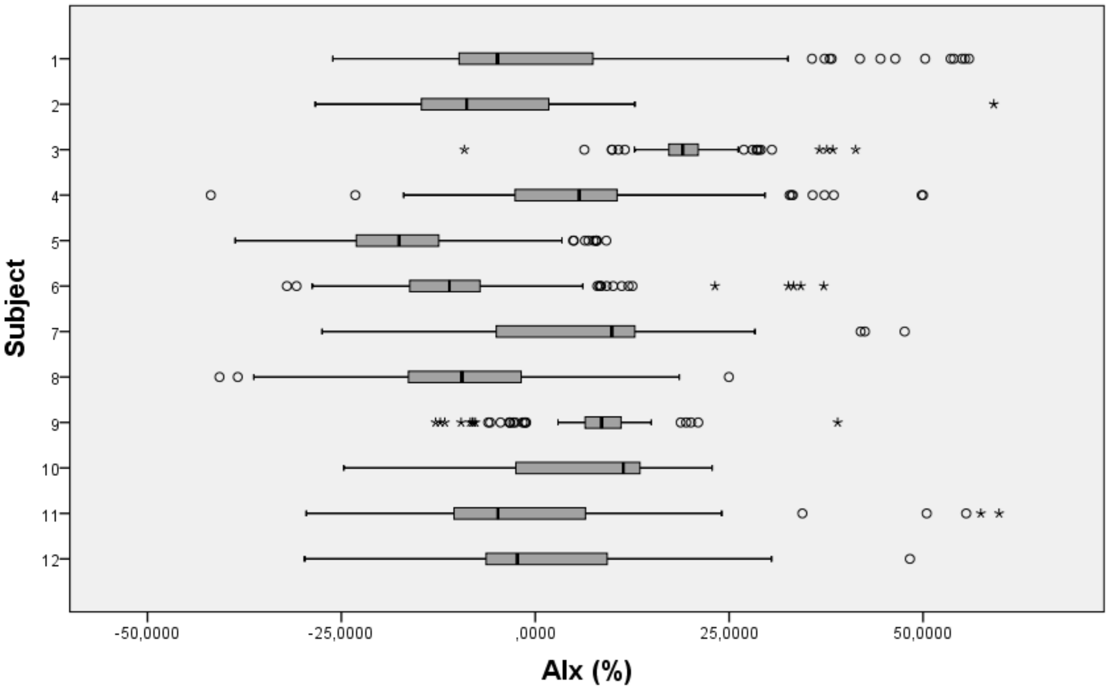


Figure 7.6 – Boxplot of the AIx categorized by subject;

Figure 7.6 is very useful for visualizing the variability on each subject. Seven subjects (1, 2, 5, 6, 8, 11 and 12) present a negative AIx as a median value, and the other five volunteers (3, 4, 7, 9 and 10) present positive AIx. As hypothesized, relevant differences exist between subjects and a particular analysis is the most correct approach.

Table 7.7 – Table of each subject AIx IQR;

AIx (%)	***Subject***											
	1	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***	***11***	***12***
IQR	17.25	*16.57*	*3.78*	*13.25*	*10.67*	*9.24*	*17.89*	*14.55*	*4.75*	*15.99*	*16.95*	*16.96*

Without any sub – categorization, it can be assessed that in there are subjects with low to medium-high variability. Subjects with medium high – variability contain a larger IQR and more extended ± 1.5 × IQR limits.

The lowest variability is found on subject 3, which also has the highest AIx of all the subjects. This can be explained by the fact that the subject is a woman and is the eldest of the sample (29 years). Women tend to have higher AIx values than men, and with aging, AIx also increases. Subject 6 and 9 also have low variability (but the latter presents excessive outliers).

Medium variability is found on subject 2, 5 and 10. A latter sub – categorization of each subject by month, carotid site and trial can be crucial in finding out where the variability truly lies.

Significant differences exist between subjects and, therefore, different ADW profiles can be found on this dataset.

7.2.3.3.1 *Sub – categorization by month*

A statistical analysis of each subject sub – categorized by each month can be helpful in assessing the differences between months on each subject. A Mann – Whitney U test was conducted to assess if these AIx dissimilarities do exist between months, in each subject. Results are expressed in Table 7.8.

Table 7.8 – Table of each subject AIx divided by month. AIx median values are indicated. Mean rank, U and P represent the conclusions given by the Mann Whitney test. A P – value of < .05 was considered as significant;

Subject	*Month*	*Median AIx (%)*	*Mean rank*	*U*	*P*
1	**Month 1**	-6.06	98.98	**-0.323**	**0.75**
	Month 2	4.48	101.65		
2	**Month 1**	-9.14	120.29	-2.345	< .05
	Month 2	-8.35	142.22		
3	**Month 1**	19.85	141.42	-3.116	< .05
	Month 2	18.54	112.67		
4	**Month 1**	6.13	75.87	**-1.076**	**0.28**
	Month 2	3.97	83.73		
5	**Month 1**	-19.68	183.45	-4.460	< .05
	Month 2	-15.76	236.21		
6	**Month 1**	-8.91	150.75	-7.637	< .05
	Month 2	-16.07	82.07		
7	**Month 1**	9.47	155.06	-3.548	< .05
	Month 2	11.08	193.66		
8	**Month 1**	-12.10	124.86	-5.577	< .05
	Month 2	-4.06	181.27		
9	**Month 1**	6.34	59.12	-3.420	< .05
	Month 2	9.26	83.70		
10	**Month 1**	-2.39	105.88	-4.812	< .05
	Month 2	11.55	150.42		
11	**Month 1**	-9.02	92.00	-6.218	< .05
	Month 2	4.62	148.38		
12	**Month 1**	13.11	117.39	-5.406	< .05
	Month 2	-3.41	74.00		

Table 7.8 indicates that only subject 1 and 4 did not present significant differences between months. Even subjects with few differences in the median (subjects 2, 3 and 5) revealed high dissimilarities between each month. This could be due to differences in the variability of each month measurement, which is directly influencing the ranks.

Figure 7.7 display AIx boxplots for each subject, divided by month. Most of the subjects present unexpected variabilities in the quartiles, between months. The given result directly agrees with the conclusion drawn from sub – section 7.2.3.1: **there are significant differences in AIx between months, as only 2 of 12 subjects presented no AIx dissimilarities in each month.**

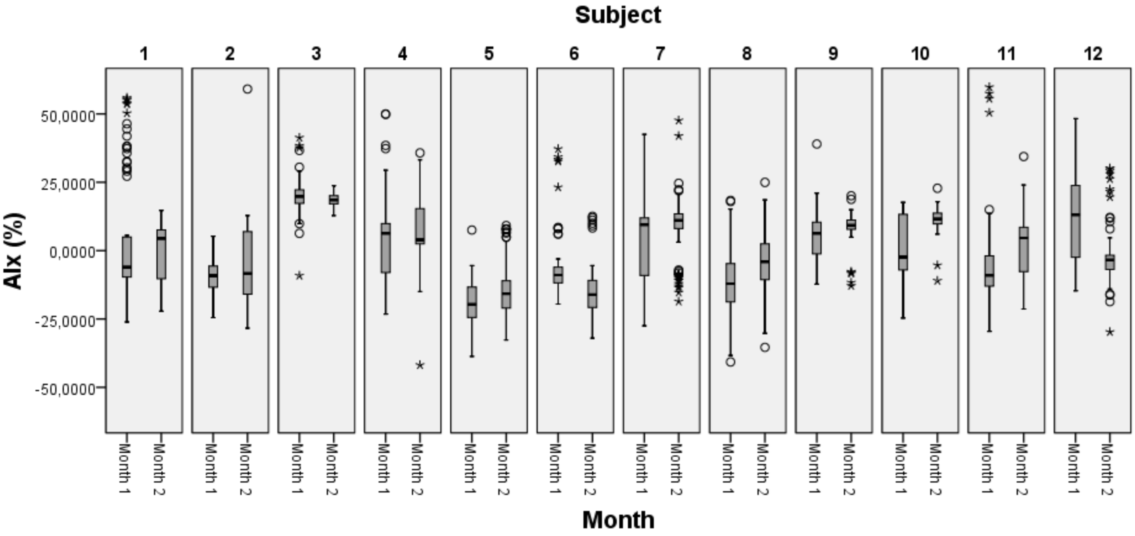


Figure 7.7 – Boxplot of the AIx categorized by subject and sub – categorized by month;

7.2.3.3.2 *Sub – categorization by carotid site*

A statistical analysis of each subject sub – categorized by each month can be helpful in assessing the differences between months on each subject. Figure 7.8 displays AIx boxplots for each subject, divided by carotid site.

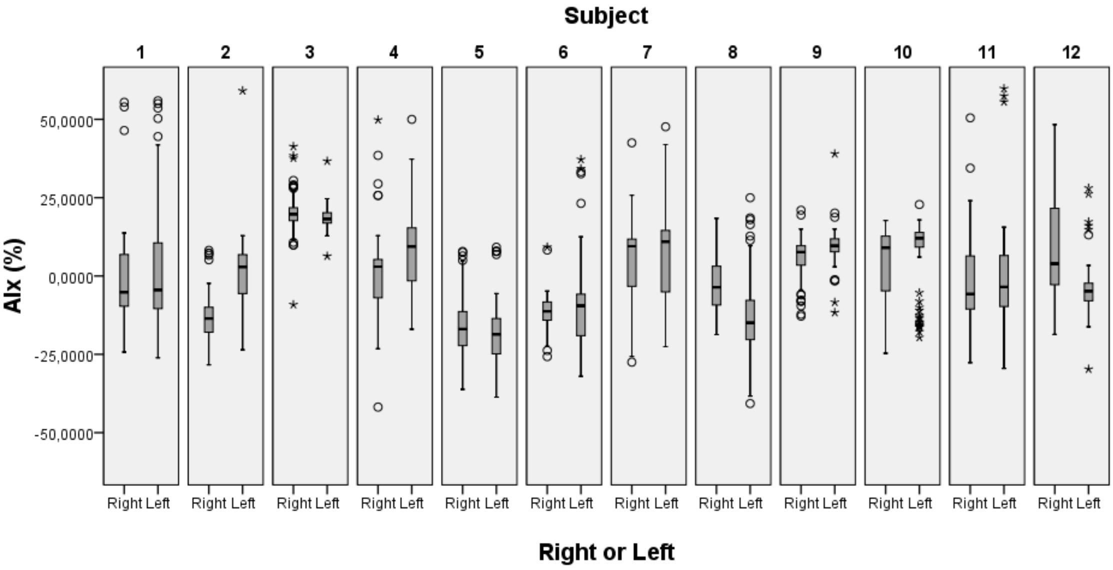


Figure 7.8 – Boxplot of the AIx categorized by subject and sub – categorized by carotid site;

While performing a visual investigation, and in most of the subjects, few differences can be spotted between the median AIx in the carotid site. Discarding the variability of left and right measurements, few or no dissimilarities between each median are found in subjects 1, 3, 5, 6, 7, 10 and 11. Subjects 4, 9 and 12 present medium median variations, while subjects 2 and 8 are highly influenced by the carotid site. However, it is important to confirm if few differences do exist with a statistical test, in this case, and once again, the best option in the Mann – Whitney test.

Table 7.9 displays the medians sub – categorized by carotid site, for each subject. A Mann – Whitney U test was performed to assess the existence of AIx dissimilarities between left and right carotid, on each subject.

Table 7.9 – Table of each subject AIx divided by carotid site. AIx median values are indicated. Mean rank, U and P columns represent the conclusions given by the Mann Whitney test. A P – value of < .05 was considered as significant;

Subject	*Carotid site*	*AIx median (%)*	*Mean rank*	*U*	*P*
1	**Right**	-5.18	98.68	**-0,502**	**0.62**
	Left	-4.46	102.82		
2	**Right**	-13.75	82.00	-11,955	< .05
	Left	2.86	194.59		
3	**Right**	19.73	139.99	**-1,822**	**0.07**
	Left	18.18	129.48		
4	**Right**	2.97	60.53	-4,769	< .05
	Left	9.42	95.38		
5	**Right**	-16.93	220.09	-2,604	< .05
	Left	-18.63	189.33		
6	**Right**	-11.28	116.63	**-0,764**	**0.45**
	Left	-9.49	123.46		
7	**Right**	9.48	157.42	-2,777	< .05
	Left	10.96	187.26		
8	**Right**	-3.62	194.37	-8,474	< .05
	Left	-14.92	109.83		
9	**Right**	7.65	60.47	-3,383	< .05
	Left	9.64	84.43		
10	**Right**	9.02	107.89	-4,216	< .05
	Left	12.04	146.97		
11	**Right**	-5.74	120.74	**-0,580**	**0.56**
	Left	-3.52	126.39		
12	**Right**	3.95	112.46	-6,308	< .05
	Left	-4.86	63.43		

Mann – Whitney U test revealed no significant differences between left and right carotid site in subjects 1, 3, 6, and 11. Unexpectedly, significant differences were found between left and right in subject 5, 7, and 10. Subject 5 revealed significant differences at P =

.05, but does not reveal significant differences at a P = .01 level (U = – 2.777, P = 0.01), which can be a clear indication that the difference is not so significant at all.

7.2.3.3.3 Sub – categorization by month and carotid site

A Mann – Whitney test is conducted to assess differences between left and right carotid in each subject, while considering each month separately. Results are quite interesting (as they are extensive, they are presented in Appendix A). For subjects 1, 3, 5, 7, 9, 11 and 12, at least one month revealed no differences in the AIx values between left and right carotid. However, subject 6 revealed significant differences in each month, which clearly contradicts the result expressed on table 7.8. This means that differences do exist in subject 6 as well.

These results and the result expressed in 7.2.3.3.2 do not contradict the conclusions that were reached in 7.2.3.2, because in most of the subjects, at least one measurement revealed that there are no significant differences between left and right carotid. However, it is enough to distrust them.

7.2.3.3.4 Sub – categorization by month, carotid site and trial

It is important to evaluate the intra – subject AIx variability. A low variability between different trials is expected, considering that the subject and the room remain in the same conditions (section 6.1) between each trial.

In each subject, trials were further sub – categorized by month and carotid site. This was done because results prior to this study revealed that there were significant differences between months, and that similarities between left and right carotid site do seem to exist, but cannot be totally trusted. Due to its power to assess if dissimilarities exist between three or more groups, the Kruskal – Wallis H test was performed between the three trials, divided by left and right carotid site, on each month. A P – value < .05 was considered as statistically significant. The table results are very extensive, and so, they are presented in Appendix B. The conclusions obtained regarding intra – subject variability follows:

- **Subject 1 –** In month 1, the three trials of each carotid site did not present significant dissimilarities [Right: X^2 = 0.491, P = 0.78, Left: X^2 = 0.539, P = 0.76]. In month 2, trials did not present significant differences only in right carotid site [X^2 = 4.798, P = 0.09].
- **Subject 2 –** In month 1, no significant discrepancies between trials was obtained only in left carotid [X^2 = 4.798, P = 0.09]. In month 2, both carotid sites revealed no significant differences between trials. [Right: X^2 = 2.606, P = 0.27, Left: X^2 = 3.915, P = 0.14].

- **Subject 3 –** In month 1, both carotid sites revealed significant differences between trials. In month 2, only right carotid site did not present dissimilarities [X^2 = 2.815, P = 0.24].
- **Subject 4 –** In month 1, only the left carotid site did not reveal significant differences in AIx between trials [X^2 = 5.664, P = 0.06]. In month 2, both carotid sites did not present significant discrepancies between trials [Right: X^2 = 0.201, P = 0.90, Left: X^2 = 1.370, P = 0.50].
- **Subject 5 –** The only non – significant result is in month 1, in the right carotid site. [X^2 = 4.426, P = 0.10].
- **Subject 6 –** In month 1, both carotid sites did not present significant differences [Right: X^2 = 3.322, P = 0.19, Left: X^2 = 0.079, P = 0.96]. In month 2, right carotid site did not present significant differences at P = .05 [X^2 = 4.503, P = 0.10]. Left carotid site presented significant differences at P = .05, but is considered as a not significant result at P = .01 [X^2 = 6.981, P = 0.03].
- **Subject 7 –** In month 1, dissimilar results between trials were obtained in both carotid sites. On the other hand, month 2 left and right carotid site trials presented results with no dissimilarities between trials. [Right: X^2 = 4.243, P = 0.12, Left: X^2 = 2.894, P = 0.24].
- **Subject 8 –** In month 1, both right and left carotid site did not have significant discrepancies between trials [Right: X^2 = 3.134, P = 0.21, Left: X^2 = 1.087, P = 0.58]. In month 2, left carotid did not present significant dissimilarities [X^2 = 1.175, P = 0.56]. Right carotid site presented significant differences between trials at P = .05, but is considered as a not significant result at P = .01 [X^2 = 6.138, P = 0.046].
- **Subject 9 –** In both months, left carotid site did not reveal significant differences in AIx between trials [Month 1: X^2 = 0.542, P = 0.76, Month 2: X^2 = 0.542, P = 0.76].
- **Subject 10 –** In month 1, left and right carotid site revealed significant differences between trials. In month 2, both carotid sites revealed no dissimilarities between trials [Right: X^2 = 0.875, P = 0.64, Left: X^2 = 5.900, P = 0.052].
- **Subject 11 –** In month 1, only the left carotid site presented no significant differences in AIx between trials [X^2 = 0.626, P = 0.73]. In month 2, the left carotid site did present no significant differences between trials, as well [X^2 = 1.264, P = 0.53].

- **Subject 12 –** In month 1, both carotid sites revealed no significant discrepancies between trials [Right: X^2 = 5.638, P = 0.06, Left: X^2 = 0.356, P = 0.84]. In month 2, the result was equally satisfactory, with both carotid sites revealing no significant differences in AIx between trials [Right: X^2 = 2.023, P = 0.36, Left: X^2 = 5.172, P = 0.08].

All subjects revealed at least one comparison between three trials that did not present significant differences. No dissimilarities between trials in both right and left carotid site in a single month can be noticed in most of the subjects (subjects 1, 2, 4, 6, 7, 8, 10, 12). In four possible not – significant outcomes for each subject, six subjects presented three not – significant results. One subject revealed four not – significant outcomes. **In conclusion, the AIx variability between trials has shown excellent results.**

7.2.4 Agreement assessment

This sub - section will reveal the results of the Bland – Altman plots. Further sub – sections are made for each of agreements of interest (see section 6.4.1.4):

7.2.4.1 Test A - Month 1 & month 2

Figure 7.9 displays a Bland – Altman AIx plot of two measurements: month 1 and month 2. The bias is -1.5%, and the upper and lower limits are approximately 11.43% and -14.43%, respectively.

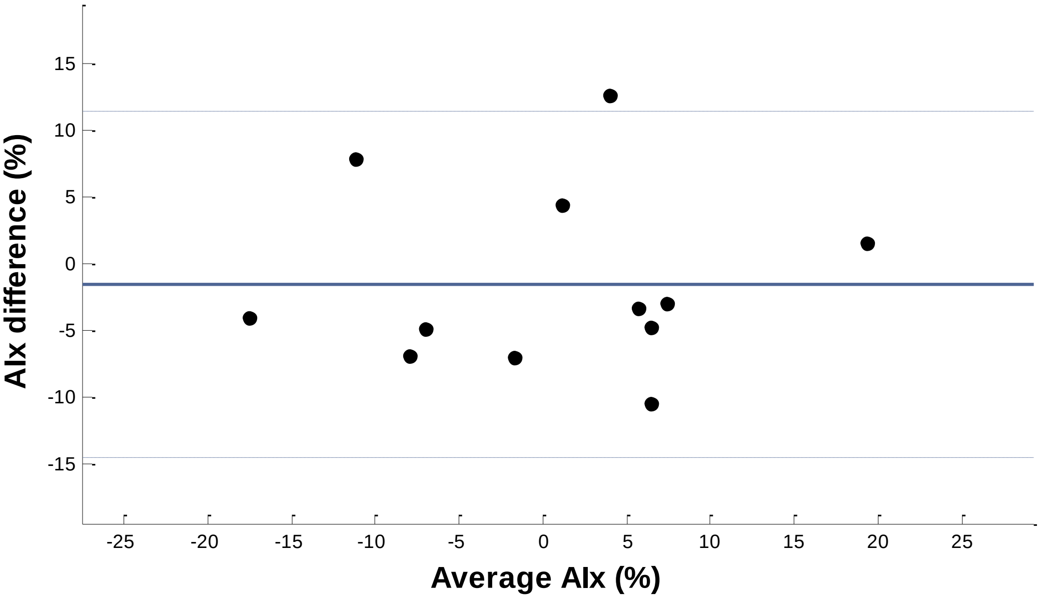


Figure 7.9 – Bland – Altman AIx plot between month 1 and month 2. The bias is -1.5%, and the upper and lower limits are 11.43% and -14.43%, respectively;

The AIx bias between month 1 and month 2 differs a bit from zero. However, it is a small difference (1.5%), and considering that we are measuring AIx and a variation of 1.5% can be considered as irrelevant, it is not clinically significant. No trends can be visualized, and the variability looks quite consistent along the graphic. Not all of differences between measurements between month 1 and month 2 lie within the limits of agreement, as 91.7% of the differences are inside the limits. This value is quite high, nonetheless, **which can be a clear indication that there is a relative agreement in AIx between months.**

7.2.4.2 Test B - Month 1 right carotid & month 1 left carotid

The following figure displays a Bland – Altman AIx plot of two measurements: month 1 right carotid and month 1 left carotid. The bias is 3.44%, and the upper and lower limits are 21.32% and -14.44%, respectively.

The AIx bias between month 1 and month 2 differs from zero by 3.44%. No trends can be visualized, but the variability increases in the extremes. All of differences between measurements between month 1 and month 2 lie within the limits of agreement.

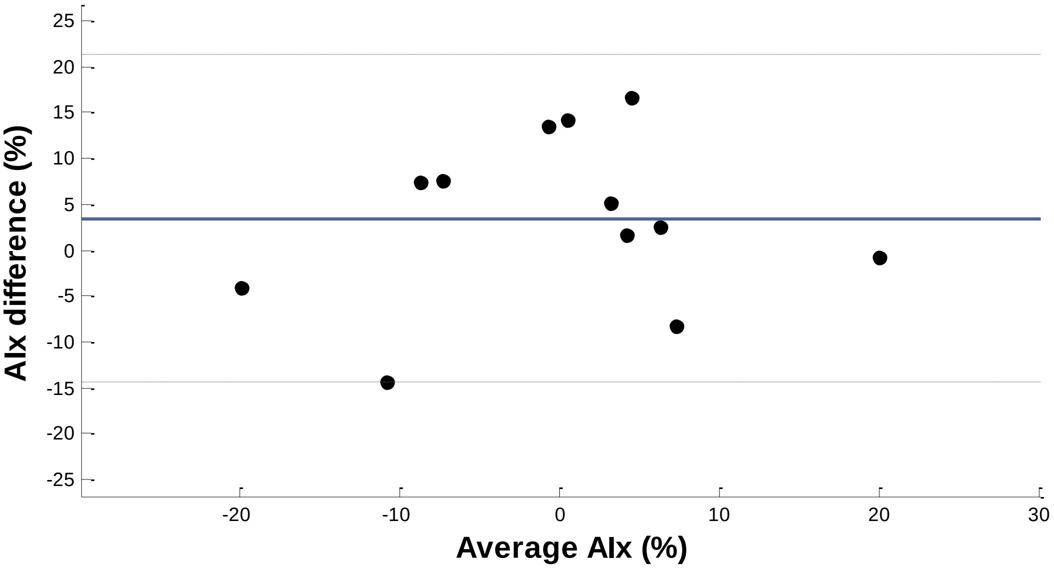


Figure 7.10 – Bland – Altman AIx plot between month 1 left carotid and month 1 right carotid. The bias is 3.44%, and the upper and lower limits are 21.32% and -14.44%, respectively;

7.2.4.3 Test C - Month 2 right carotid & month 2 left carotid

Figure 7.11 displays a Bland – Altman AIx plot of two measurements: month 2 right carotid and month 2 left carotid. The bias is -0.70%, and the upper and lower limits are 16.03% and -17.44%, respectively.

The bias is relatively low (0.70%), which means that there is low discrepancy between left and right measurements in month 2. Variability is consistent, and no patterns can be visualized. Only one of the differences between measurements does not lie within the limits of agreement, which represents relatively good agreement. A comparison between test B bias with the given bias for test C reveals that month 1 measurement was of lower quality while compared with month 2. However, **there is enough agreement between AIx values in left and right carotid site on each month.**

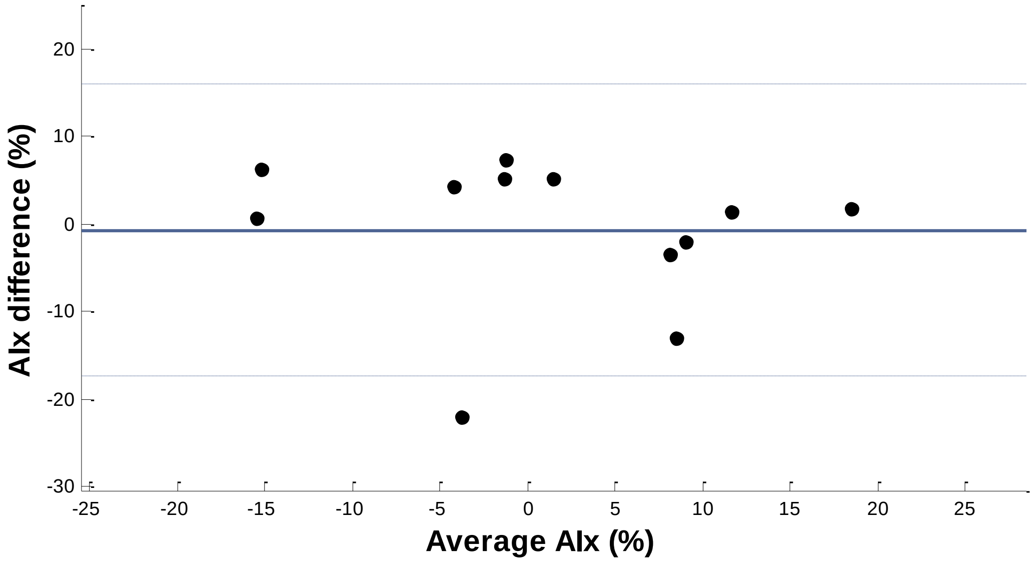


Figure 7.11 – Bland – Altman AIx plot between month 2 left carotid and month 2 right carotid. The bias is -0.70%, and the upper and lower limits are 16.03% and -17.44%, respectively;

7.2.5 Dataset I – Discussion

There are significant differences in AIx between months, as only 2 of 12 subjects presented no AIx dissimilarities in each month (sections 7.2.3.1 and 7.2.3.3.1).

A Bland – Altman AIx plot revealed low bias and relative agreement between months (sections 7.2.4.1, 7.2.4.2 and 7.2.4.3).

The fact that there are significant differences in AIx between months does not necessarily mean there is lack of agreement between months. A low bias revealed that both measurements revealed very similar results on average. There is relative agreement, as 91.7% of the differences between months are inside the limits of agreement. But, in spite of that, the Bland – Altman method assumes that 95% of measurements must be inside both limits of agreement, so there is not a total agreement.

Some problems must have interfered with the measurements while assessing month – to – month repeatability. As revealed in 7.2.4.2 and 7.2.4.3, month 1 revealed a bias that was higher than unexpected, which could mean that the quality of the measurement was not as good as in the second month. Also, misplacement of the sensor at the carotid or differences in the subject/room that passed unnoticed to the operator while doing the second measurement can be some of the possible explanations. More studies need to be performed between sessions and/or months.

No AIx dissimilarities are found not only between measurements in right and left carotid in each month, **but also between left and right carotid without month categorization** (section 7.2.3.2)**. Bland – Altman plots have also revealed that there is enough agreement between AIx values in left and right carotid site on each month** (sections 7.2.4.2 and 7.2.4.3)**.**

A further statistical analysis categorized by subject revealed that those dissimilarities were not so significant at all, as for most of the subjects, at least one measurement revealed that there are significant differences between left and right carotid. Results were not totally contradicted, and the assumption that there are no differences between left and right is maintained. However, these results could be mistrusted (section 7.2.3.3.2 and 7.2.3.3.3)**.**

The subject by subject analysis revealed that more studies are needed to ultimately confirm that there are no differences between left and right carotid in a young dataset. Still, there is enough evidence to assume that few differences seem to exist between left and right carotid site AIx in dataset I, which is composed of young subjects (18 – 30 years). Still, one should always consider the possibility (even if improbable) that a cardiovascular disorder in one of the carotid arteries of a subject may be interfering with given results. Also, for some subjects, one of the carotid sites was harder to capture efficiently, which could definitely influence the acquired ADW and consequently, the extracted parameter values.

The fact that few differences exist between left and right coincides with a recent study from Luo *et al.* (2011) [80]. This study concluded that dissimilarities in left and right carotid only start to exist from 35 years onwards, with the left carotid becoming thicker than right carotid due to hemodynamic and biochemical different effects on the carotid intima – media thickness of each carotid site. Increased AIx is associated with increased intima – media thickness, as both are associated with CV risk. As the dataset I age range is between 18 – 30 years, no differences between left and right carotid site should be observed. And results from dataset I indicate that in both months, the left carotid seems to be slightly thicker (due to higher AIx), but not enough to be considered as a statistically significant result.

A more rigid approach in 'catching' the carotid to acquire the ADW of each subject and wider time measurements (≈ 60 – 80 seconds) are important guidelines for improvements in future carotid site measurements. A 'live' ADW quality control procedure right after the acquisition method could be very useful, so that each of the signals that are stored in the database have assured top – quality.

The AIx variability between trials has proved excellent results, with good AIx repeatability between trials (section 7.2.3.3.4)**.**

The fact that a good AIx repeatability between trials was ascertained is extremely important, as it strengthens the possibility that differences between months and left and right carotid site could be due to sensor displacement. During each of the three trials, sensor is not displaced, and remains fixed in the same position, and considering that the seated subjects remains seated, comfortable and in the same temperature controlled room, it is expected that the results should correspond to good repeatability.

While practicing scientific investigation, the quality of the research is directly related with the quality of the used instruments. Therefore, this result also assures that the uncertainties regarding repeatability between months and between left and right carotid are not due to hardware, firmware underperformance, software imperfections, but mainly due to overlooked guidelines in the acquisition process between months. More rigid and demanding patient measurement protocols could and should be implemented, to minimize even further any possible external influences.

7.3 Dataset II – Angioplasty: Case Study

Figure 7.12 shows a set of three – four pulses detected invasively and non – invasively, before and after carotid intervention.

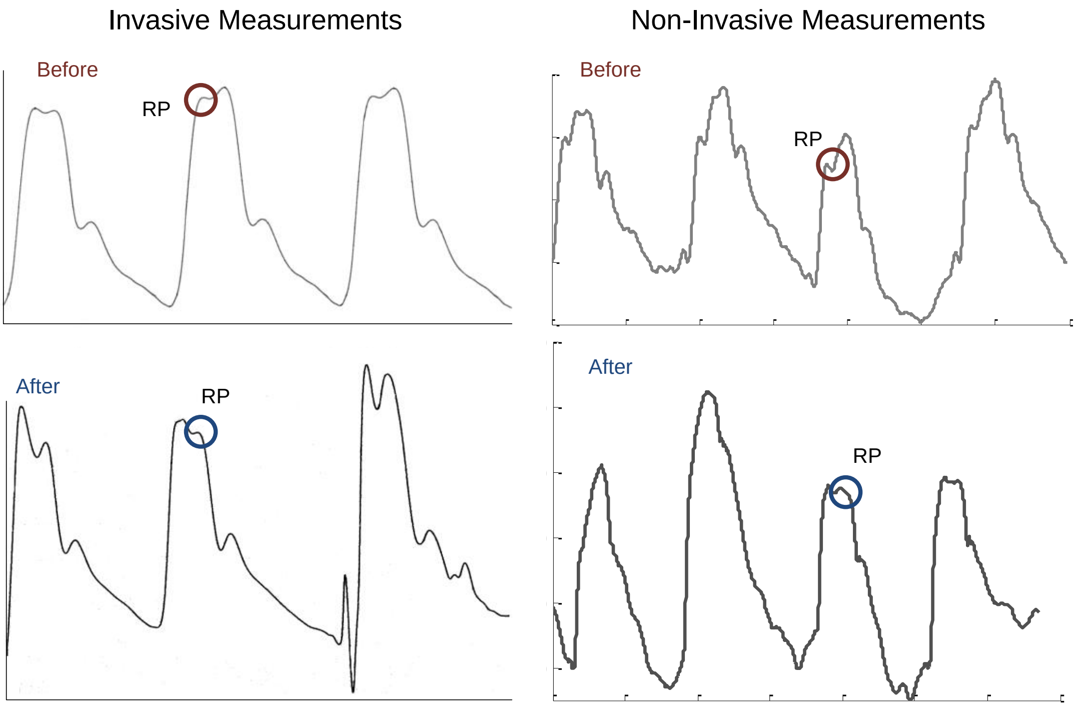


Figure 7.12 – Set of three – four pulses that were detected invasively and non – invasively, before and after carotid intervention, for the same subject. RP_T for a specific pulse is represented by red and blue circles, corresponding to APW collected before and after carotid intervention, respectively;

ADW physiological differences can be denoted before and after angioplasty, revealing that anomalous reflection waveform disappeared with the carotid intervention. This phenomenon is detected not only by the reference method, but also by the developed non – invasive system.

The SP_T, RP_T and DN_T attributes were measured for the non – invasive method before and after the angioplasty procedure (Figure 7.13). As expected, DN_T associated time did not change, due to the inexistence of known cardiac valves complications. However, visible changes occurred in the SP_T and RP_T analysis. Before the surgical procedure, RP_T occurs earlier than SP_T, but after the intervention, this tendency is inverted, with SP_T occurring first than RP_T.

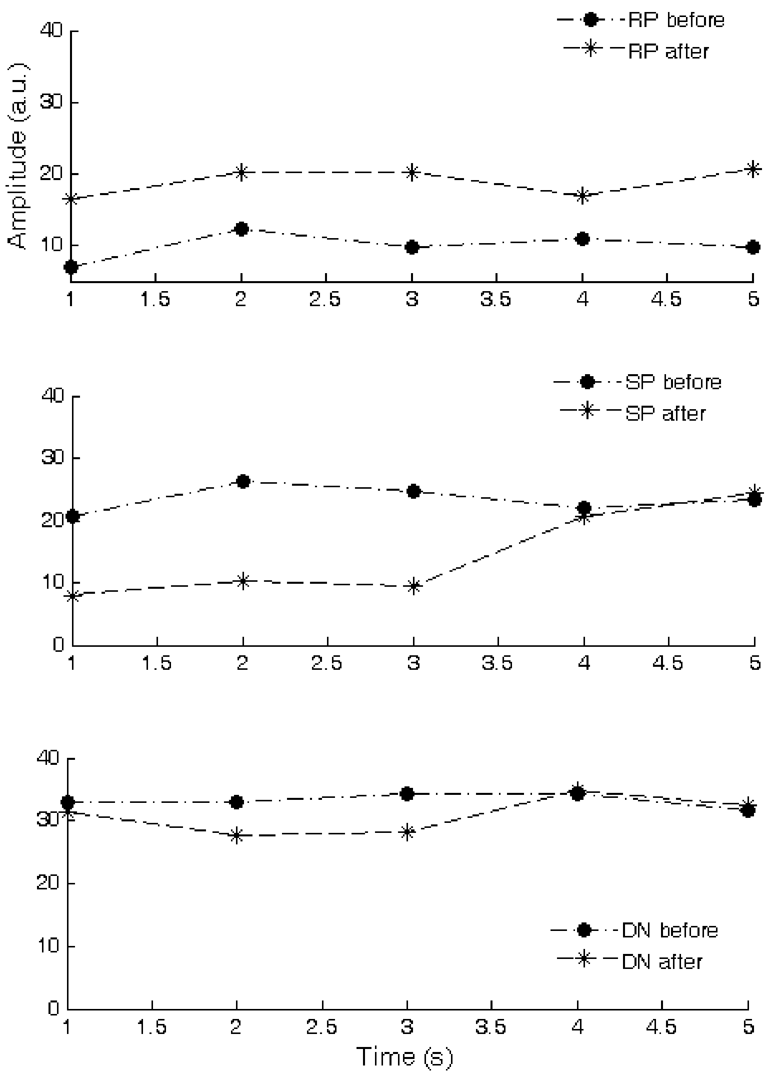


Figure 7.13 – SP$_T$, RP$_T$ and DN$_T$ time parameterization comparison before and after carotid intervention;

7.3.1 Dataset II – Discussion

In general, non – invasive measurements are less accurate when compared with invasive trials, mainly due to the dependency of conditions of measurement and operator. However, the non – invasive PZ probe is just as capable of effectively detecting physiological modifications before and after angioplasty procedures. This proves the usefulness of this technology, and can facilitate early identification of cardiac problems through screening trials.

7.4 Dataset III – Classification

Neural network classification methods were the first to be constructed, as a performance study for MLP (both with 1 and 2 hidden layers) and RBF involving a meticulous parameter tuning was arranged. Usually, the neural network with higher accuracy was chosen, but sensitivity, specificity and precision measures also play an important role, and were not neglected.

7.4.1 Dataset III characterization

Dataset III histogram characterization for each attribute follows in Figure 7.14. Blue histograms are associated with sub – group I (healthy subjects), while red histograms are associated with sub – group II (unhealthy subjects). These histograms display the class distribution in each attribute in a small matrix, as the full descriptive analysis is extensive. A full overview of the data is presented in Appendix C.

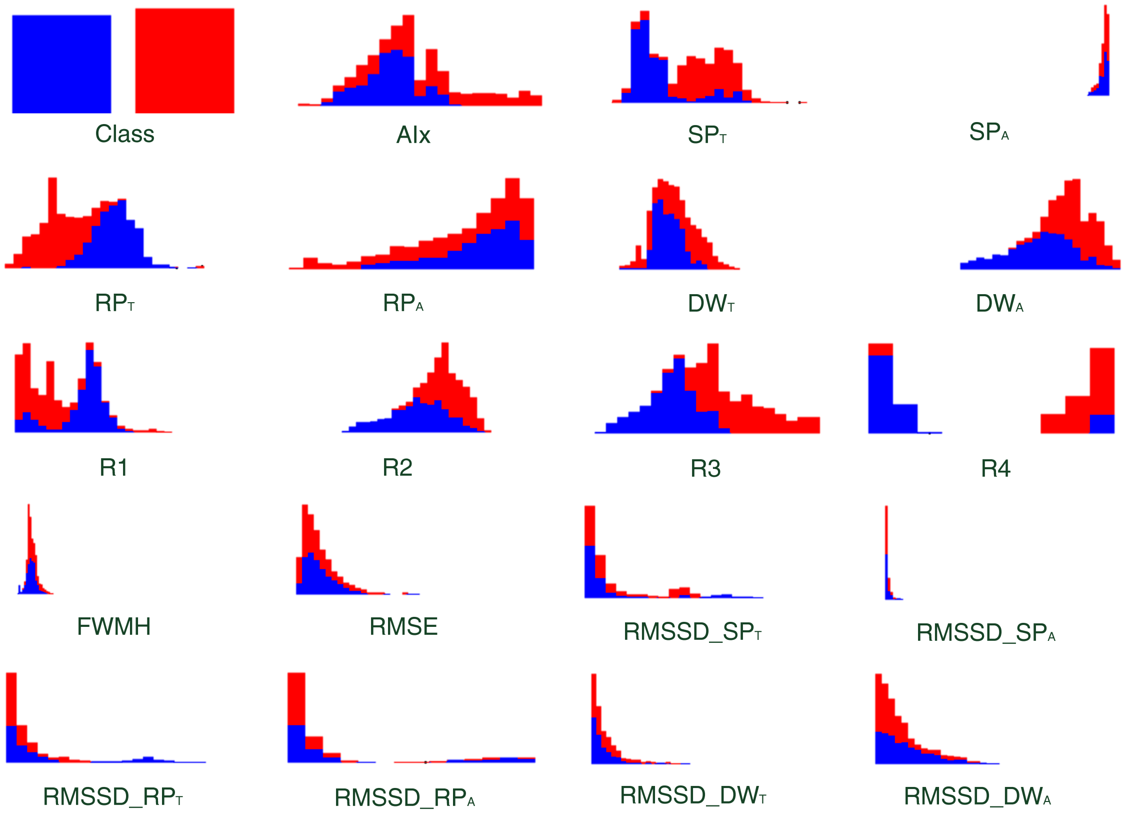


Figure 7.14 – Dataset III characterization. Blue histogram represents sub – group I (healthy) and red histogram is associated to sub – group II (unhealthy);

Analyzing the prominent points that were extracted from the ADW, it can be assessed that RP_T occurs earlier than SP_T for sub – group II, while sub – group displays the opposite, as expected. The RP_T histogram displays sub – group II predominance in lower values, and the SP_T histogram shows sub – group I prevalence in lower SP_T values. The AIx histograms display a predominance of the healthy sub - group in the left part of the total histogram, as expected. SP_A and RP_A do not seem do hide any potentially useful information regarding dissimilarities in arterial stiffness function. DN_T and DN_A do not seem to reveal any important particularity, as well. However, red histograms have a slight prevalence in higher DN_T and DN_A values.

Ratios R1 – R4 present important visual information, as there is a clear formation of 2 sub – groups in each ratio histogram, and can be important attributes in procedures that involve group formation from raw data (clustering, for example). As for other attributes, RMSSD attributes do not have any visual dissimilarity. RMSE does not seem to be a significant attribute for discerning between the two sub – groups, as well.

7.4.2 ANN performance study

Different parameter configurations were tested for each neural network, while manipulating one of the parameters and comparing the accuracy, sensitivity, specificity and precision values.

7.4.2.1 MLP (1 – hidden layer)

Different runs were performed while manipulating one single parameter. Training method, hidden neurons, training time (in epochs), learning rate (%) were adjusted and prematurely, only accuracy was measured. The default run (which is, in fact, the default Weka configuration for the MLP) is indicated in Table 7.10 as the grey row, and may be used as a 'baseline' model for direct performance comparison between configurations.

Some interesting facts can be observed in Table 7.10. Changing the *k* value in *k* – fold CVN doesn't seem to maximize accuracy when compared to the default configuration (run 2 - 4). A changing to a higher training time is useless (run 10) and modifying the learning rate doesn't seem the change accuracy as well (run 11 – 14). However, the manipulation of the number of hidden neurons is an important issue. When more hidden neurons are added, the accuracy seems to increase (Run 6 and 7).

Table 7.10 – 1 – hidden layer MLP performance study, while manipulating diverse MLP characteristic parameters. The grey row indicates the default Weka configuration for 1 – hidden layer MLP;

Run	***Training method***	***Hidden neurons***	***Training time (epochs)***	***Learning rate (a. u.)***	***Accuracy (%)***
1	10 – fold CVN	10	500	0,3	96.64
2	5 – fold CVN	10	500	0,3	96.06
3	20 – fold CVN	10	500	0,3	96.30
4	40 – fold CVN	10	500	0,3	96.51
5	10 – fold CVN	0	500	0,3	94.03
6	10 – fold CVN	20	500	0,3	**96.98**
7	10 – fold CVN	40	500	0,3	96.70
8	10 – fold CVN	10	1	0,3	87.78
9	10 – fold CVN	10	100	0,3	94.64
10	10 – fold CVN	10	1000	0,3	96.53
11	10 – fold CVN	10	500	0,1	96.53
12	10 – fold CVN	10	500	0,01	93.52
13	10 – fold CVN	10	500	0,2	96.33
14	10 – fold CVN	10	500	0,5	96.33

As the ‘hidden neurons’ parameter seems to have capital importance, it was manipulated a little bit further. Comparing run 6 and run 7, the first presents itself as having higher accuracy, which suggests that the optimal number of hidden neurons may be approximately 20. While maintaining the default configuration for the other parameters, the ‘hidden neurons’ parameter was continuously changed for values between 15 and 25. Table 7.11 expresses these results, while adding sensitivity, specificity and precision measures to assess which configuration is the best.

Table 7.11 shows that the optimal value might be between 19 and 20 neurons, for run 5 and 6, respectively. Each run has the same accuracy, with run 5 having higher sensitivity and precision. On the other hand, run 6 has higher specificity. Each of the runs can be used as a model for future classifier comparison. Run 6 is chosen because of its higher specificity, which is quite important considering that healthy and unhealthy subjects are being studied. One should consider models with higher specificity, because it is preferable having a classifier with few FPs as possible, to avoid misdiagnosing unhealthy patients as healthy. However, methods with high sensitivity shouldn't be discarded, as high FNs rates can also lead to other problems regarding patient physical and psychological discomfort.

Table 7.11 – 1 – hidden layer MLP performance study, while manipulating the 'hidden neurons' parameter. The grey row indicates the default Weka configuration for 1 – hidden layer MLP;

Run	*Hidden neurons*	*Accuracy (%)*	*Sensitivity (%)*	*Specificity (%)*	*Precision (%)*
1	15	96.06	95.43	96.64	96.38
2	16	96.57	96.56	96.58	96.35
3	17	96.67	96.56	96.78	96.56
4	18	96.84	96.63	97.04	96.85
5	19	**96.98**	**96.84**	97.10	**96.91**
6	20	**96.98**	9677	**97.17**	96.83
7	21	96.70	96.07	97.10	96.40
8	22	96.77	96.98	96.58	96.37
9	23	96.34	95.93	96.71	96.47
10	24	96.60	96.35	96.85	96.62
11	25	96.34	96.35	96.32	96.08

Therefore, the preferred configuration for the 1 – hidden layer MLP is the default configuration with the 'hidden neurons' parameter equal to 20.

7.4.2.2 2 – hidden layers MLP

Similarly to 7.4.2.1, different runs were performed while manipulating one single parameter. Training method, hidden neurons, training time and learning rate were the parameters of choice. Accuracy, sensitivity, specificity and precision measures helped in finding the best model, in overall. The default run is indicated in Table 7.12 as the grey row.

Learning rate changes did not impress once again in 2 – hidden layers MLP. However, this time, changes in the training method revealed some interesting results. Run 3 gives the best values of specificity and precision from all runs overall. However, the sensitivity value is quite lower than one should expect. On the other hand, run 9 revealed the best value of sensitivity, but neglects specificity. Run 11 presents itself as a more balanced run, with higher accuracy as well. **And so, run 11 is the chosen run for future classifier comparison.**

Table 7.12 – 2 – hidden layers MLP performance study, while manipulating diverse MLP characteristic parameters. The grey row indicates the default Weka configuration for 2 – hidden layers MLP;

Run	***Training Method***	***Hidden Neurons***	***Learning Rate***	***Accuracy (%)***	***Sensitivity (%)***	***Specificity (%)***	***Precision (%)***
1	10 – fold CVN	10, 10	0.3	96.44	96.49	96.33	96.15
2	5 – fold CVN	10, 10	0.3	95.89	95.23	96.52	96.24
3	20 – fold CVN	10, 10	0.3	96.64	*95.79*	**97.44**	**97.22**
4	40 – fold CVN	10, 10	0,.3	96.61	96.21	96.98	96.75
5	10 – fold CVN	10, 10	0.01	93.96	94.25	93.69	93.33
6	10 – fold CVN	10, 10	0.1	95.79	96.77	94.88	94.65
7	10 – fold CVN	20, 20	0.3	96.51	96.56	96.45	96.22
8	10 – fold CVN	5, 20	0.3	96.30	95.93	96.65	96.40
9	10 – fold CVN	20,5	0.3	96,.13	**97.12**	*95.20*	94.99
10	10 – fold CVN	21,21	0.3	96.44	96.42	96.45	96.22
11	10 – fold CVN	22,22	0.3	**96.74**	96.52	96.91	96.70

7.4.2.3 RBF

Two parameters were varied in the RBF performance study: clustering seed (a random estimate of the initial weight) and number of clusters (which is, in fact, the number of RBF). Results are expressed in Table 7.13.

It is possible to note that the accuracy, sensitivity, specificity and precision of the default configuration, when compared with the MLP, is much lower. However, by manipulating the parameters, it was possible to obtain similar MLP accuracies. Clustering seed is not a definitive factor in raising accuracy, but run 3 was the best while modifying the clustering seed value. The number of clusters was of crucial importance, with run 9 presenting the best results. Therefore, in runs 11 – 13, we used the same number of clusters as run 9 (100) and manipulated the clustering seed for 3 values: 1, 5, and 10.

Table 7.13 – RBF performance study, while manipulating diverse MLP characteristic parameters. The first row indicates the default Weka configuration for 2 – hidden layers MLP;

Run	*Clustering seed*	*Number of clusters*	*Accuracy (%)*	*Sensitivity (%)*	*Specificity (%)*	*Precision (%)*
1	1	2	91.92	92.56	91.33	91.92
2	5	2	92.74	94.32	91.26	92.74
3	*10*	*2*	*92.81*	*94.11*	*91.59*	*92.81*
4	20	2	92.74	93.61	91.92	92.74
5	1	5	92.26	90.46	93.96	92.26
6	1	10	94.49	94.25	94.71	94.49
7	1	20	95.32	95.02	95.60	95.32
8	1	50	95.79	94.46	97.04	95.79
9	*1*	*100*	*96.03*	*94.74*	***97.24***	*96.03*
10	1	250	95.32	94.53	96.06	95.32
11	5	100	**96.23**	**96.23**	95.79	**96.65**
12	1	100	**96.23**	95.23	97.17	**96.23**
13	10	100	95.76	95.12	96.36	96.13

Run 11 and run 12 present the best results, with the same accuracy (96,23%), the same precision (96,23%), but different values of sensitivity and specificity, with run 11 being more sensitive and run 12 being more specific. Before, in 7.4.2.1 and 7.4.2.2, runs with higher accuracy and specificity were preferred. This time, **run 11 was the opted run, as RBF has as it is the most sensitive RBF model.**

7.4.3 Classifier selection

After the neural network performance study, the other classifiers were constructed using 10 – fold CVN. Accuracy, sensitivity, specificity and precision were measured, and RMSE values were computed. Training speed was also measured, and training speed ranges were discretized into 5 groups, even though it was not a factor in choosing the three best classifiers. Training speed guidelines are shown in Table 7.14.

Table 7.14 – Discretized guidelines for different training speed ranges;

Range (s)	*Discretized value*
0 – 2	*****
2 – 10	****
10 – 20	***
20 – 60	**
> 60	*

Classifier selection results follow up in Table 7.15, with the column with ranking by descendent accuracy included.

Table 7.15 – Classifier selection results, with rankings assessed by descendent accuracy;

Classifier	*Training speed (s)*	*RMSE (%)*	*Accuracy (%)*	*Sensitivity (%)*	*Specificity (%)*	*Precision (%)*	*Ranking*
MLP (1 HL)	**	16.09	96.98	96.77	97.17	96.83	**2**
MLP (2 HL)	*	17.46	96.74	96.56	96.91	96.70	**3**
RBF	***	18.58	96.23	96.23	95.79	96.65	4
C4.5	****	20.10	95.72	95.58	95.86	95.58	5
Random forest	****	14.60	97.15	97.47	96.85	96.66	**1**
RIPPER	***	19.89	95.72	96.14	95.34	95.07	6
Naïve Bayes	*****	34.00	88.05	86.60	89.42	88.46	8
Bayesian network	*****	32.16	89.01	84.84	92.90	88.46	7

All classification methods except Bayesian based - classification exhibited accuracy values > 95%. Random forest was the best classifier, in overall, by having the highest accuracy and sensitivity. Comparing with the other classifiers, MLP ANN methods require high computational resources, as they can take between 30 seconds and several minutes to be trained. ANN have shown excelling results in the specificity department, which means that the FP error rate is inferior, in other words, misdiagnosing unhealthy patients as healthy is less likely to happen when using MLP classifiers.

In the global classifier performance study above, the three best classifiers with highest accuracy were selected: random forest, 1 – hidden layer MLP and 2 - hidden layers MLP.

7.4.4 Diagnostic with 3 – best multiple classifier methodology

The three best classifiers were combined and applied to sub – group III, which is a group of 10 undiagnosed subjects to predict their CV condition. The predicted class depends on the percentage of positive and negative pulses (1 = healthy and 2 = high CV risk factor, respectively) for each classifier. An output diagnosis is then obtained combining the three classifier results, based on an equal – weighted voting system. Results are shown in Table 7.16.

Table 7.16 – Diagnostic with a 3 – best multiple classifier methodology;

	Random forest			*MLP (1 – HL)*			*MLP (2 – HL)*			*Final Result*	
Subject							*% Pulses*				
	1	*2*	*Result*	*1*	*2*	*Result*	*1*	*2*	*Result*	*Ratio*	*Decision*
1	64	36	1	85	15	1	92	8	1	3/3	1
2	12	88	2	72	28	1	72	28	1	2/3	1
3	2	98	2	93	7	1	33	67	2	1/3	2
4	0	100	2	77	23	1	12	88	2	1/3	2
5	70	30	1	98	2	1	87	13	1	3/3	1
6	89	11	1	99	1	1	99	1	1	3/3	1
7	0	100	2	33	67	2	20	80	2	0/3	2
8	53	47	1	91	9	1	36	64	2	2/3	1
9	12	88	2	76	24	1	62	38	1	2/3	1
10	0	100	2	77	23	1	13	87	2	1/3	2

The assumption that the classification of an undiagnosed subject should not be based on the result of one, and only one classifier is supported by the given results, as the global final result was never the same final result for each classifier:

- In subject 2, for example, the random forest model defines the subject as having high CVD risk. However, both MLP output is contradictory, and due to that, the final decision is that the subject is healthy.

- Other examples are subject 3, 4 and 10, where random forest and 2 – hidden layers MLP agreed that the subject could have a high CVD risk, while 1 – hidden layer MLP assessed these subjects as healthy.
- Subject 8 is an example where 2 – hidden layers MLP defined the subject as having high CVD risk. Although, the final result is that the subject is healthy, as random forest and 1 – hidden layer MLP defined him as a healthy patient.

There are also subjects with explicit full agreement between the three classifiers (subjects 1, 5, 6 and 7). Even though we obtained high accuracy values for each of the used classifier models, at least once there was a contradiction between a single model and the other two models.

The results of this multiple classifier diagnostic procedure clearly show the versatility and usefulness of this method. By using an ensemble model like the proposed, the error rate is minimized and accurate diagnostics are more probable and feasible.

7.4.5 Dataset III - Discussion

In the global classifier performance study, the three best classifiers that were selected for the subsequent diagnostic methodology were: random forest, 1 – hidden layer MLP and 2 - hidden layers MLP (section 7.4.3)**.**

The fact that random forest was the best classifier in terms of accuracy confirms the effectiveness and superior performance of the classifier, which was already suggested in a previous study with a similar dataset by Almeida *et al.* (2011) [26].

3 ANN trained classifiers were ranked in the top 4 out of 8 possible classifiers. This demonstrates that when ANN problems regarding low tolerance to noise and bad comprehensibility are attenuated and/or bypassed, coupled with studies to maximize the accuracy of the model, ANN have superior classification performance, as hypothesized by Kotsiantis (2007) [67]. MLP – based ANN have also shown excelling results in the specificity department, which means that the FP error rate is low, in other words, misdiagnosing unhealthy patients as healthy is less likely to happen when using MLP classifiers.

The results of this multiple classifier diagnostic procedure clearly show the versatility and usefulness of the method. By using an ensemble model like the proposed, the error rate is minimized and accurate diagnostics are more probable and feasible (section 7.4.4)**.**

The implementation of a model that is based on the guidelines of the work by Gorunescu *et al.* (2011) [32] successfully demonstrated that the assumption that the classification of an undiagnosed subject should not be based on the result of one and only one classifier.

7.5 Dataset IV – Clustering

Clustering methods are very useful when little or no information is known from a dataset. Consequently, risk group assessment in a healthy population became a definite opportunity with this dataset.

7.5.1 Dataset IV characterization

Dataset IV characterization follows in Figure 7.15. More detailed figures (with mean ± SD included) are annexed in Appendix D.

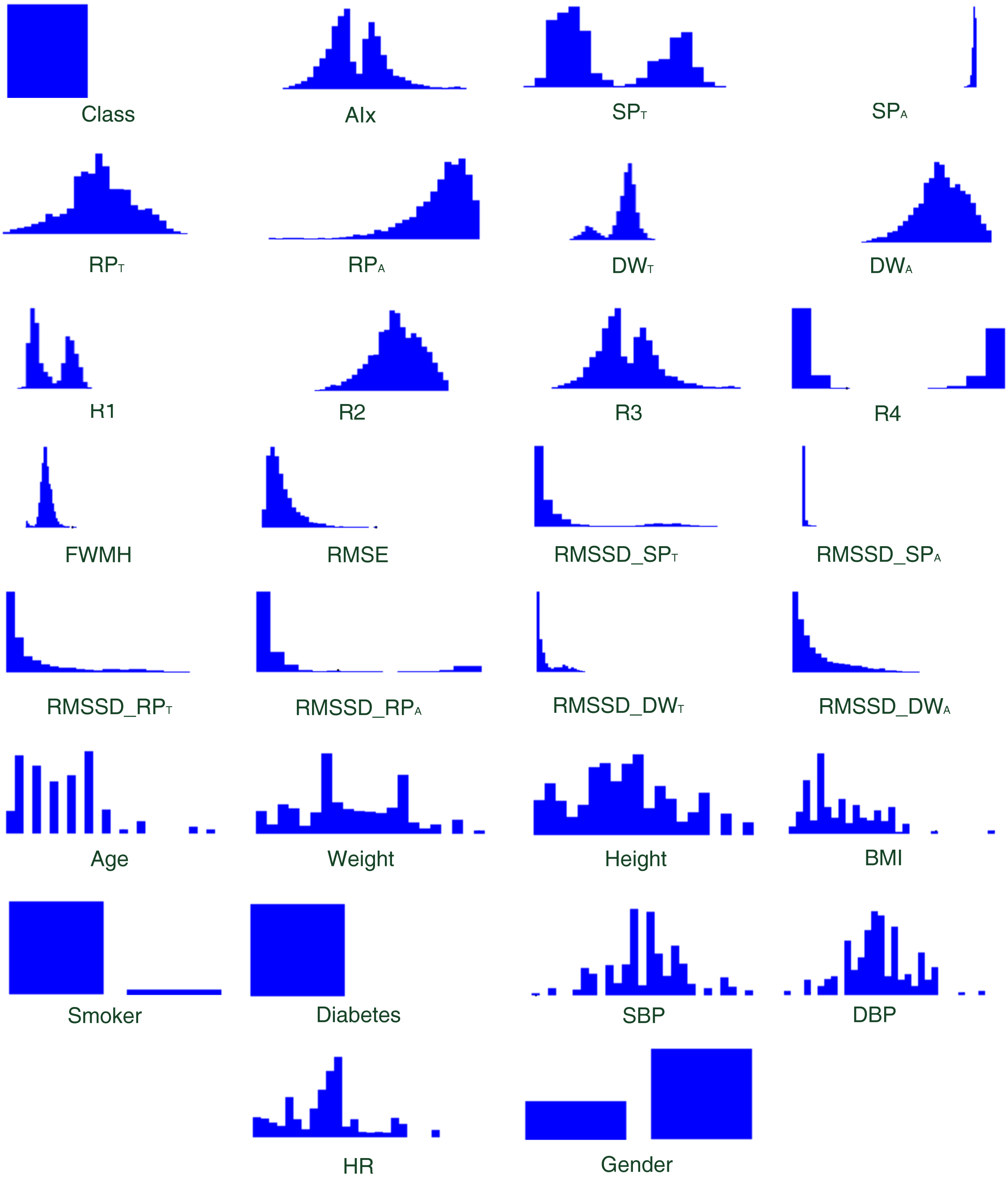


Figure 7.15 – Dataset IV histogram characterization;

A primary visual histogram analysis to attributes AIx, SP_T, R1, R3 and R4 suggests the existence of two or more possible sub – groups.

7.5.2 Two risk group clustering

Figure 7.16 displays both plots of the RP_T as a function of SP_T after 2 - class EM (a) and *k* – means clustering (b). Different colored points represent different pulse labels (blue = class 1; red = class 2). Categorical features were eliminated in the *k* – means clustering method, as this technique has some serious versatility problems with categorical attributes.

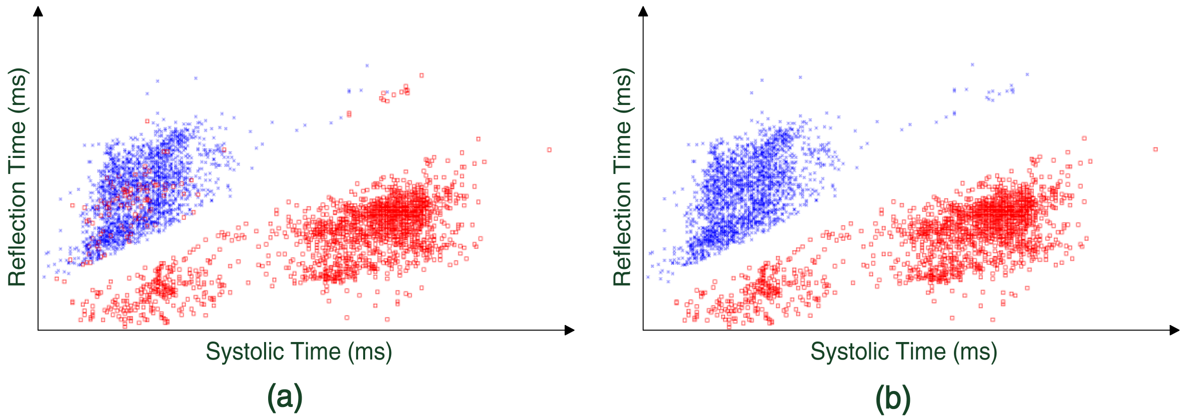


Figure 7.16 – Plots of RP_T as a function of SP_T after clustering in (a) EM and (b) *k*-means clustering for two risk group assessment. Blue = class 1 and red = class 2;

The EM plot presents visually average results, as some pulses of class 1 look completely mislabeled. Visually, the results of *k* – means clustering are more desirable and satisfying, as it partitioned the dataset in two perfectly homogeneous risk groups are possible with this dataset. For both EM and k – means, the blue group (class 1) presents itself as a more healthier partition than the red group (class 2), as it represents the points where the reflection wave arrived after the systolic point, and so, it can be assumed that class 1 represent a group with a lower risk of developing CVD and class 2 represents a group with higher CVD risk. However, the fact that the reflection point is after the systolic point does safely means that the subject is healthy, as some other attributes may influence the pulse class determination.

When analyzing for differences in the cluster centroids, between clustering methods (Table 7.16), some small dissimilarities can be found. The biggest differences are found in SP_T, AIx and R4. It is possible that the k – means algorithm considers these three attributes with higher average merit than the EM algorithm. An attribute subset selection for both algorithms can be performed to erase these doubts (see section 7.4.1.1).

Another visual interpretation of Table 7.17 is needed to find out where are the biggest differences between the class labels 1 and 2. At first glance, the biggest differences lay in AIx, SP$_T$, RP$_T$, R3 and R4. RMSSD_SP$_T$ and RMSSD_RP$_T$ results are also quite dissimilar. Another important conclusion is that the group with higher risk (EM - cluster 2) has approximately 85% of pulses from women gender. **This might suggest that in the dataset IV age range, women have higher AIx than men.**

Table 7.17 – Results for both clustering methods in two risk group assessment. Class 1 and class 2 represent the blue and red clusters of figure 7.15, respectively. Each scalar attribute in each cluster is the cluster centroid (mean value for each cluster);

Attribute		*Full Data*	*Clustering Method*			
			K - means		*EM*	
			1	*2*	*1*	*2*
Cluster pulses		4471	2463	2008	2315	2156
Gender	***M***	1327	--	--	**993**	**334**
	F	3144			**1322**	**1822**
Smoker	***Yes***	231	--	--	108	123
	No	4240			2207	2033
Diabetes	***Yes***	0	--	--	--	--
	No	4471				
Age		21.41	21.66	21.11	21.71	21.11
Weight		60.46	63.21	57.09	63.17	57.55
Height		1.67	1.70	1.65	1.70	1.65
BMI		21.36	21.75	20.88	21.73	20.96
SBP		107.89	108.62	106.99	108.67	107.05
DBP		69.48	68.90	70.17	68.85	70.15
HR		71.01	72.88	68.71	73.05	68.81
SP$_T$		**163.45**	**117.06**	**220.34**	**116.64**	**213.67**
RP$_T$		**159.29**	**179.87**	**134.04**	**179.96**	**137.10**
DW$_T$		281.99	274.36	291.34	273.72	290.86
SP$_A$		0.996	0.9965	0.9954	0.9966	0.9954
RP$_A$		0.87	0.89	0.85	0.89	0.86
DW$_A$		0.74	0.70	0.80	0.69	0.80
R1		**118.54**	**157.29**	**71.00**	**157.07**	**77.19**
R2		0.75	0.70	0.81	0.70	0.80
R3		**0.0033**	**-0.11**	**0.14**	**-0.11**	**0.12**
R4		**-0.11**	**-0.89**	**0.86**	**-0.89**	**0.74**
AIx		**0.35**	**-11.23**	**14.55**	**-11.38**	**12.93**
RMSSD_SP$_T$		24.35	21.68	27.62	18.43	30.69
RMSSD_SP$_A$		0.0036	0.0029	0.0044	0.0011	0.0063
RMSSD_RP$_T$		41.60	49.27	32.19	49.30	33.33
RMSSD_RP$_A$		0.16	0.18	0.14	0.17	0.15
RMSSD_DW$_T$		35.51	42.77	26.61	43.53	26.91
RMSSD_DW$_A$		0.09	0.11	0.06	0.11	43.52
FWHM		448.21	433.43	466.34	431.39	466.26
RMSE		0.0529	0.0540	0.0515	0.0536	0.0521

7.5.2.1 Attribute subset selection

The gain ratio attribute evaluator results for both clustering methods are expressed in the graphic of Figure 7.17, which shows the average merit of each attribute in determining each pulse class label.

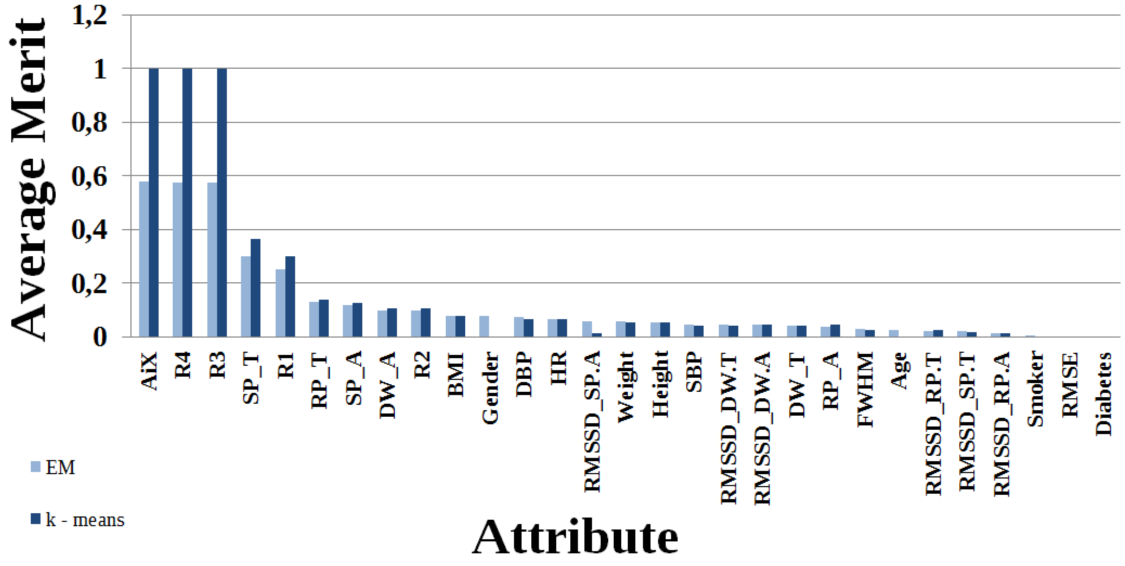


Figure 7.17 – Attribute subset selection for each clustering method in two risk group assessment;

As expected, AIx is the most important attribute in defining class labels. All ratios (except R2) and wave reflections (SP_T and RP_T) are also important in distinct group definition. As for demographic attributes, BMI and gender have the highest average merit values, although they don’t look significantly important. As suggested before, by visual plot analysis, age does not seem to be an important factor. Other attributes like smoker, RMSE and diabetes seems to be irrelevant.

The successful clustering in two risk groups suggests that AIx and wave reflections phenomenon have a crucial importance in CV risk assessment in young subjects. Surprisingly, age does not seem to be an important factor in determining risk groups for subjects between 18 and 30 years.

7.5.3 Three risk group clustering

The *k* – means clustering was the chosen algorithm for three risk group clustering, as the observed results for k – means were visually better than EM. Just like in two risk group clustering, categorical features were eliminated, as k – means has some serious incompatibilities with categorical attributes. The clustering for three risk group assessment is presented in Figure 7.18.

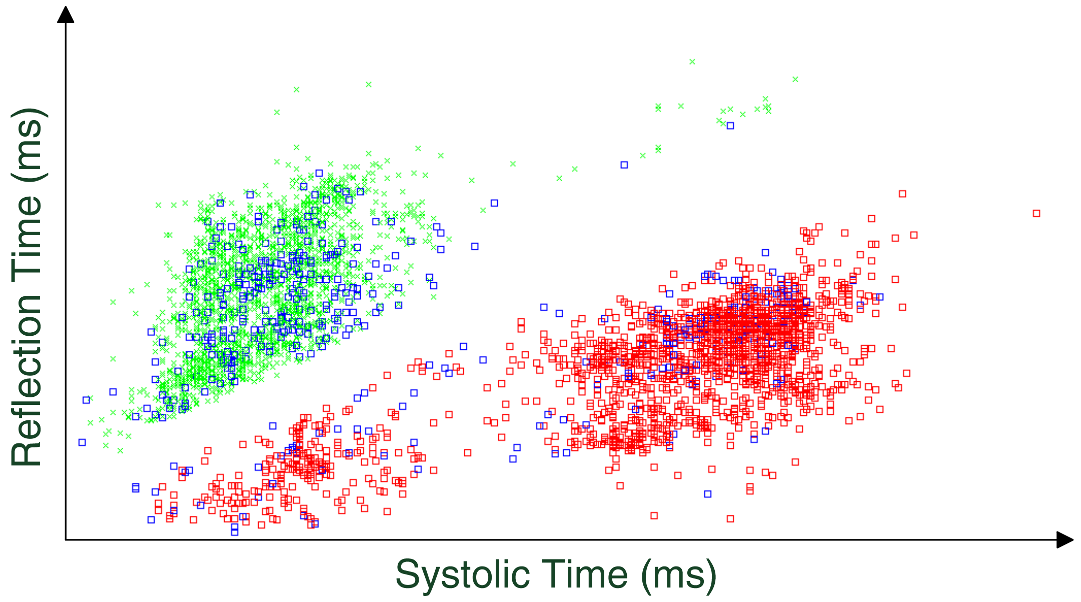


Figure 7.18 – Plot of RP_T as a function of SP_T after *k*-means clustering for three risk group determination. Green = class 1, blue = class 2 and red = class 3;

As observed in Figure 7.18, a complete visual distinction between three risk groups could not be materialized. This could be due to irrelevant features taking part in the clustering process, and/or due to low robustness of the algorithm. A progressive attribute removal could be performed to attempt a visual correction. However, Table 7.17 does reveal a whole different perspective.

Analyzing Table 7.18 and correlating with Figure 7.18, it can be acknowledged that cluster 1 (green homogeneous zone in figure 7.17) is mostly represented by ADW type C pulses, where $RP_T > SP_T$. Adding to that, the mean AIx is negative (-11.49), which represents cluster 1 as a low CVD risk group. Cluster 3 (red points in Figure 7.18) pulses are mostly type A and type B pulses, where $SP_T > RP_T$. Cluster 3 could represent a group with higher CV risk. Cluster 2 pulses (blue points in Figure 7.18) represent the less homogeneous group, as they are scattered in the whole plot. They are a combination of type B and type C pulses, with the similar mean AIx of the full data. They can represent an intermediate group in terms of CV risk.

With three risk group clustering, it was possible to assess groups which represent different combinations of ADW types, which could be helpful in assessing CV risk development.

Table 7.18 – Results for both clustering methods in two risk group assessment. Cluster 1, cluster 2 and cluster 3 represent the green, blue and red clusters of figure 7.17, respectively. Each scalar attribute in each cluster is represented as mean;

Attribute	*Full Data*	*Cluster 1*	*Cluster 2*	*Cluster 3*
Number of Pulses	*4471*	*2126*	*540*	*1805*
Age	21.41	21.73	21.14	21.13
Weight	60.46	23.21	61.82	56.81
Height	1.67	1.70	1.69	1.65
BMI	21.36	21.70	21.61	20.88
SBP	107.89	108.67	108.53	106.77
DBP	69.48	69.07	68.37	70.29
HR	71.01	72.94	71.82	68.49
SP_T	**163.45**	**116.58**	**154.30**	**221.39**
RP_T	159.29	180.22	162.41	133.70
DW_T	281.99	274.41	276.39	292.59
SP_A	0.996	0.9965	0.9955	0.9957
RP_A	0.87	0.89	0.90	0.85
DW_A	0.74	0.69	0.74	0.80
R1	**118.54**	**157.83**	**122.09**	**71.20**
R2	0.75	0.70	0.75	0.81
R3	**0.0033**	**-0.11**	**-0.02**	**0.14**
R4	**-0.11**	**-0.89**	**0.23**	**0.85**
AIx	**0.35**	**-11.49**	**-1.66**	**14.90**
RMSSD_SP_T	24.35	19.92	42.14	24.23
RMSSD_SP_A	0.0036	0.0026	0.0053	0.0042
RMSSD_RP_T	**41.60**	**29.89**	**163.60**	**18.88**
RMSSD_RP_A	**0.16**	**0.063**	**0.90**	**0.055**
RMSSD_DW_T	35.51	39.97	62.35	22.23
RMSSD_DW_A	0.09	0.11	0.12	0.05
FWHM	448.21	432.25	450.56	466.31
RMSE	0.0529	0.0538	0.0577	0.0504

7.5.3.1 Attribute subset selection

Results of the gain ratio attribute evaluator with 10 – fold CVN mode for the three clusters follow on Figure 7.19.

The outcome is identical to two risk group differentiation results. R3, R4, AIx and wave reflections attributes are among the highest rated attributes, **strengthening the hypothesis that AIx and wave reflections phenomenon have a capital importance in CV risk assessment in younger subjects.** For three risk group assessment, RMSSD_RP_T and RMSSD_RP_A also assume relative importance. Demographic data continues to be quite irrelevant, with BMI being the highest rated demographic attribute. Age is once again lowly rated.

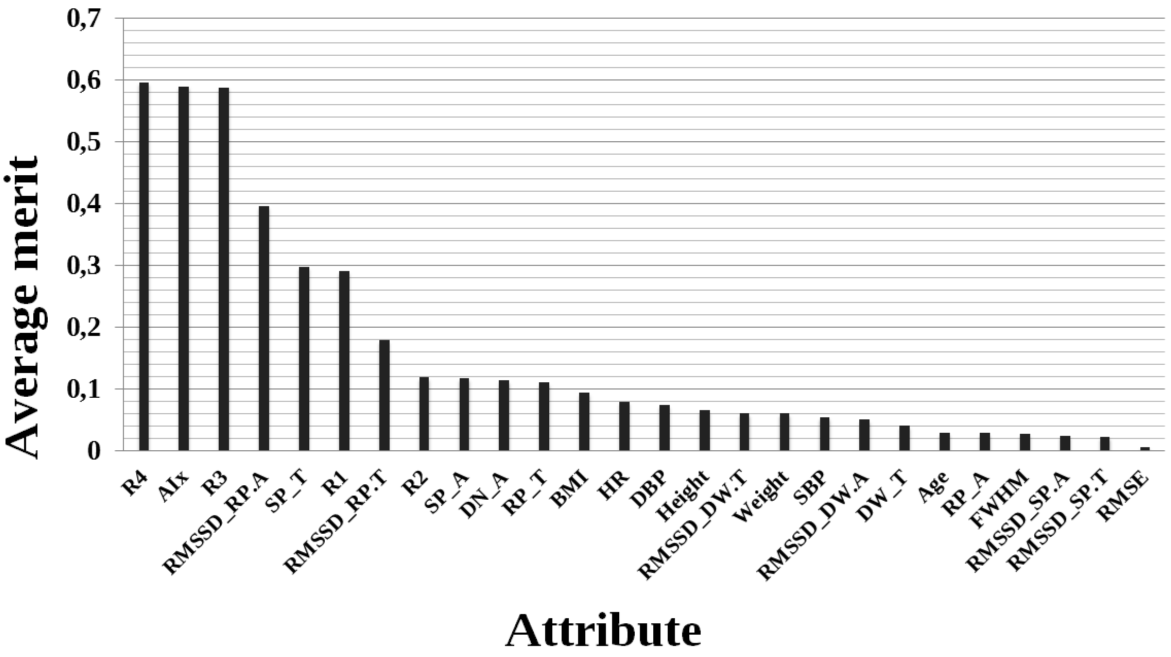


Figure 7.19 – Attribute subset selection for *k* – means clustering in three risk group assessment;

7.5.4 Dataset IV - Discussion

Two risk groups clustering results suggest that in the same age range, women have higher AIx than men (section 7.5.2)**.**

This result confirms the conclusions reached by Janner *et al.* (2010) [46] and Chung *et al.* (2010) [47], which determined that, in the same age range, women have higher AIx than men.

Age does not seem to be an important factor in determining risk groups for subjects between 18 and 30 years (section 7.5.2)**.**

Aging is the most important determinant in arterial stiffness assessment. With aging, arterial stiffness increases, and the risk of CVD is higher as well. The fact that age was not an important factor in determining different risk group is possibly due to the short age range of dataset IV. Still, it can also be an indication that an increase in the arterial stiffness due to the aging process is not so predominant until 30 years old. More studies need to be performed on this area.

The successful clustering in two risk groups suggests that the wave reflections phenomenon and carotid AIx have a crucial importance in CV risk assessment in young subjects (section 7.5.2)**.**

Data from McEniery *et al.* (2005) [49] revealed that central AIx might be a more sensitive marker of arterial aging in young and middle – age individuals (< 50 years) and aortic PWV is more sensitive in the older population (> 50 years), in healthy individuals. Although aortic PWV could not be measured in this dataset, carotid AIx (which is a direct surrogate of central AIx and wave reflections) has shown a crucial importance in determining arterial aging, and consequently, CV risk, as it was the most important attribute in terms of average merit in the partitioning process.

The success of two risk group clustering also demonstrates the efficiency of the PZ sensor in acquiring distinct and accurate CV information in a healthy population.

With three risk group clustering, it was possible to assess groups which represent different combinations of ADW types, which could be helpful in assessing CV risk development. The hypothesis that AIx and wave reflections phenomenon have a capital importance in CV risk assessment in younger subjects was also strengthened. (section 7.5.3)**.**

Excluding type D ADW, other ADW types are a clear indication of the CV condition of a subject. With the three risk group clustering, it was possible to determine three different risk groups that could be directly correlated with type A, type B and type C predominance. Identically to two risk group assessment, CV three risk group determination demonstrated that AIx and wave reflections phenomenon were the most important attributes in determining different clusters.

8. Conclusion & Future Work

After the academic project is concluded, it is necessary to evaluate and provide the main contributions that this work has given to the scientific community. Future work guidelines are also provided.

8.1 Main contributions

This work has given the following main contributions to the academic community:

1. **A significant increase in the subjects included in the database, which will undoubtedly be useful in future work.**

2. **The repeatability of the previously developed ADW acquisition system was successfully validated:**
   - The non – invasive system repeatability between months and between left/right carotid did not present the best results. There are some indications of no differences between left and right carotid for the 18-30 age range. However, the excellent repeatability results between trials, for each subject, have shown that dissimilarities are not due to hardware, firmware or software imperfections.
   - Bland – Altman plots have shown sufficient agreement between months and between left and right carotid.
   - A performed case study in subjects with stenosis also accomplished in proving that the developed PZ probe is able to detect physiological alterations after surgical procedures with good accuracy.

3. **The successful use of data mining techniques for the development of innovative decision support systems:**
   - Classification techniques have shown their usefulness in assessing CVD risk, especially ANN classifiers, which were never approach in the context of this work, and displayed superior results in comparison with other classifiers, especially in the specificity department. Also, it was demonstrated the full potential of multiple classifier methodologies in producing more trustworthy diagnostic outputs in comparison with single classifier analysis.

- Clustering procedures can be important in the premature determination of different CVD risk groups, as it was possible to assess completely distinct risk groups in a young dataset. AIx and wave reflections have proven their important in the partitioning of distinct clusters, where each cluster indicates different CV risk groups. It was also proven that age is not an important factor is risk group determination in the 18 – 30 age range, and the fact that women have higher AIx than men, in the same age range was also confirmed.
- Finally, the success of the data mining techniques is also correlated with the efficiency of the PZ sensor in acquiring accurate CV information from healthy and unhealthy subjects.

4. **A creation and development of a tool for repeatability agreement tests using the Bland Altman method, which is now open – source for future improvements. Other several optimization improvements in previously developed GUIs were also made.**

It is safe to say the main objectives proposed at the start of this work were **fully accomplished.**

8.2 Future work

The final section of this work sets new, ambitious objectives, which will surely help in defining future work methodologies. Future investigation guidelines are presented, by topics:

1. **Conclusive assessment of the non – invasive system repeatability between months and between left/right carotid.**

Results for the repeatability between months and between left/right carotid were very promising, but still inconclusive.

Some minor changes in the ADW acquisition protocol can be important in assuring few differences in measurements between months, as for example, the measured subject should refrain from tobacco and/or coffee consumption 3 hours before the measurement. As for left/right repeatability, a more rigid approach in the operator ‘catching’ the carotid and wider time measurements (≈ 60 – 80 seconds) could be important in improving data quality. However, for wider time measurements, high computational resources are needed.

2. **Reproducibility assessment of the non – invasive system.**

Reproducibility is the variability of the measurement system caused by differences in operator behavior [51]. It is important to assess it, to demonstrate if the PZ sensor is sensitive to differences in operator or not, by having two or more operators performing clinical trials, and assessing agreement between operators on each subject. At the moment, reproducibility studies are already undergoing in GEI.

3. **Repeatability and reproducibility assessment with “gold – standard" devices comparison.**

To further validate the developed non – invasive system, a comparison with results from “gold – standard” devices (SphygmoCor®, for example) is absolutely needed to evaluate and demonstrate the accuracy of the PZ sensor in providing CV information that agrees with the golden standard instrument.

4. **Database improvements for future establishment of AIx reference values for the Portuguese population.**

The database already has 155 subjects, and was significantly increased in past year, especially in the 18 – 30 age range. However, there is still a lack of data from unhealthy subjects. There are few data from healthy subjects in the 30 – 50 age range as well. The acquisition of data from the referred groups can be important in having a database with sufficient subjects to perform an establishment of AIx reference values for the Portuguese population.

5. **Implementation of biochemical attributes in future trials.**

Recent studies have determined that CVD progression is marked by the inflammatory indicator CRP and that early indicators of heart attack are the inflammatory marker CD40 and the cardiac myofilament protein troponin [81, 82]. There is a clear indication that biochemical parameters can be interesting attributes to be included in future trials, so that it is possible to assess their impact in CV risk assessment with data mining procedures. However, there is a clear difficulty in obtaining the expression of these markers.

6. **Implementation of new classification algorithms.**

After proving the usefulness of ANN in the context, other classification methods should be tested to assess their efficiency in the CVD prediction field. Support vector machines (SVMs) can be of particular interest, as they are referred in the literature as the classification technique with higher accuracy in overall, despite their problems regarding the transparency of the results and low tolerance to noise and missing values [67].

7. Clustering techniques in unhealthy patients.

Instead of performing clustering techniques in a healthy sample for the determination of different CV risk groups, there is interest in applying clustering techniques in a sample that is composed of patients with different CVD, with the objective of assessing the power of the clustering methodologies in partitioning the data into groups that represent different CV pathologies.

Appendix A – Dataset I (1)

AIx table and boxplot of each subject, categorized by carotid site and month

Table A.1 – AIx table of subjects 1-10 divided by month and carotid site. Mean rank, U and P represent the conclusions given by the Mann Whitney test. A P – value of < .05 was considered as significant;

Subject	*Month*	*Carotid site*	*AIx median (%)*	*Mean rank*	*U*	*P*
1	***Month 1***	**Right**	-7.77	40.45	**-1.896**	**0.06**
		Left	-2.60	54.16		
	Month 2	**Right**	5.78	64.13	2.372	< .05
		Left	-7.91	49.16		
2	***Month 1***	**Right**	-12.10	49.33	-7.369	< .05
		Left	-5.62	100.64		
	Month 2	**Right**	-15,57	35.02	-8.780	< .05
		Left	6.83	91.16		
3	***Month 1***	**Right**	19.99	67.14	**-0.482**	**0.63**
		Left	19.47	63.75		
	Month 2	**Right**	19.32	77.32	-4.883	< .05
		Left	17.64	45.92		
4	***Month 1***	**Right**	-5.94	32.55	-2.809	< .05
		Left	8.56	48.41		
	Month 2	**Right**	3.19	28.19	-4.249	< .05
		Left	17.28	49.63		
5	***Month 1***	**Right**	-18,32	129.35	-3.597	< .05
		Left	-21,89	96.95		
	Month 2	**Right**	-15.08	93.68	**-0.818**	**0.41**
		Left	-16.22	87.32		
6	***Month 1***	**Right**	-11.16	44.80	-5.383	< .05
		Left	-6.24	87.55		
	Month 2	**Right**	-11.74	69.41	-2.756	< .05
		Left	-19.87	37.08		
7	***Month 1***	**Right**	8.32	92.38	-2.756	< .05
		Left	10.83	115.32		
	Month 2	**Right**	10.79	65.75	**-0.929**	**0.35**
		Left	11.21	72.08		
8	***Month 1***	**Right**	-4.79	127.59	-9.291	< .05
		Left	-17.98	55.93		
	Month 2	**Right**	-2.75	68.69	-2.545	< .05
		Left	-8.03	52.51		
9	***Month 1***	**Right**	7.02	29.08	**-0.262**	**0.79**
		Left	6.14	30.30		
	Month 2	**Right**	7.75	27.88	-4.015	< .05
		Left	9.79	52.06		
10	***Month 1***	**Right**	-5.12	42.48	-5.779	< .05
		Left	12.79	80.31		
	Month 2	**Right**	12.05	73.46	-2.410	< .05
		Left	11.19	57.54		

Table A.2 – AIx table of subjects 11 and 12 divided by month and carotid site. Mean rank, U and P represent the conclusions given by the Mann Whitney test. A P – value of < .05 was considered as significant;

Subject	*Month*	*Carotid site*	*AIx median (%)*	*Mean rank*	*U*	*P*
11	***Month 1***	**Right**	-9.02	56.56	**-0.045**	**0.96**
		Left	-11.99	56.18		
	Month 2	**Right**	6.21	75.48	-3.000	< .05
		Left	-2.89	55.39		
12	***Month 1***	**Right**	15.42	35.07	**-1.672**	**0.09**
		Left	2.40	23.56		
	Month 2	**Right**	-1.23	75.45	-4.431	< .05
		Left	-4.89	46.89		

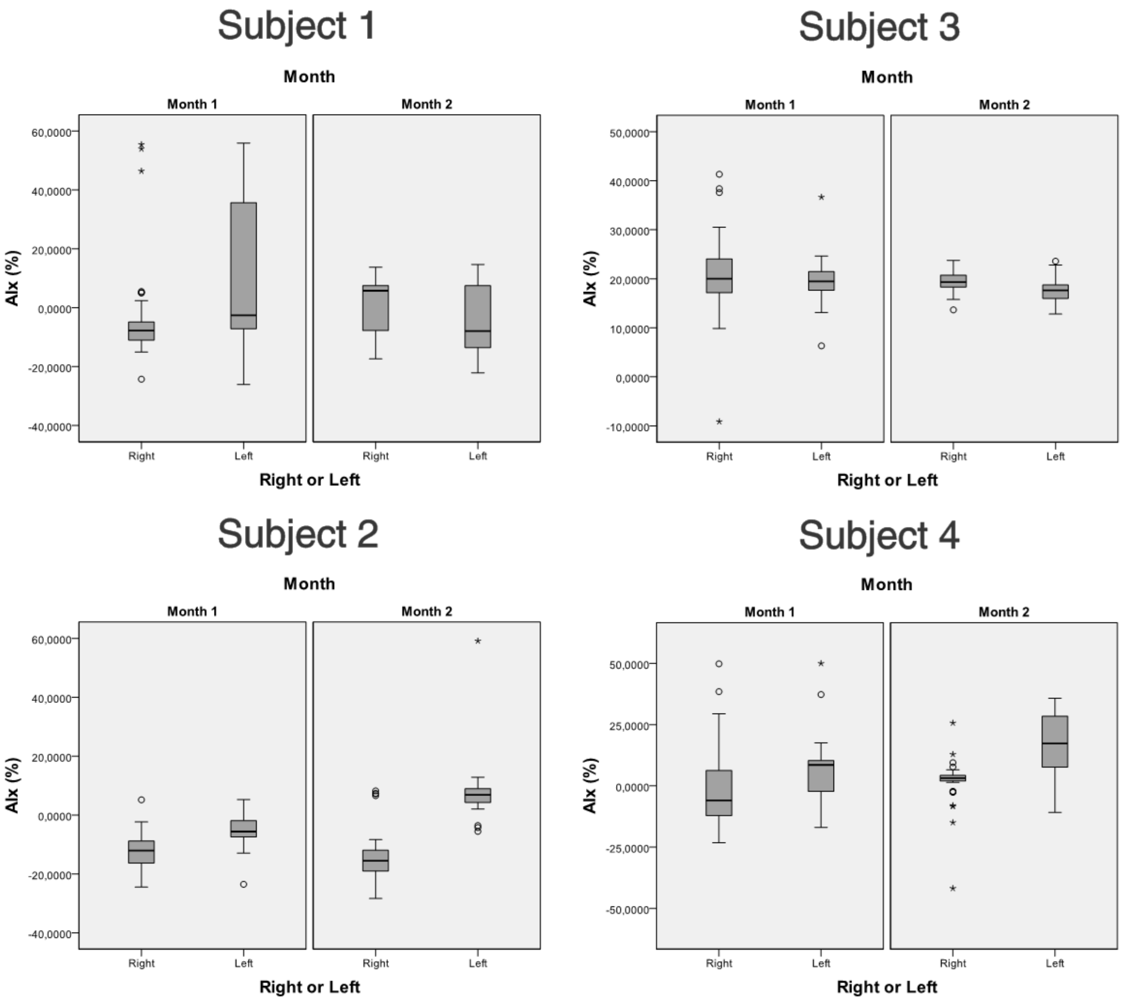


Figure A.1 – AIx boxplots of subjects 1-4 divided by month and carotid site;

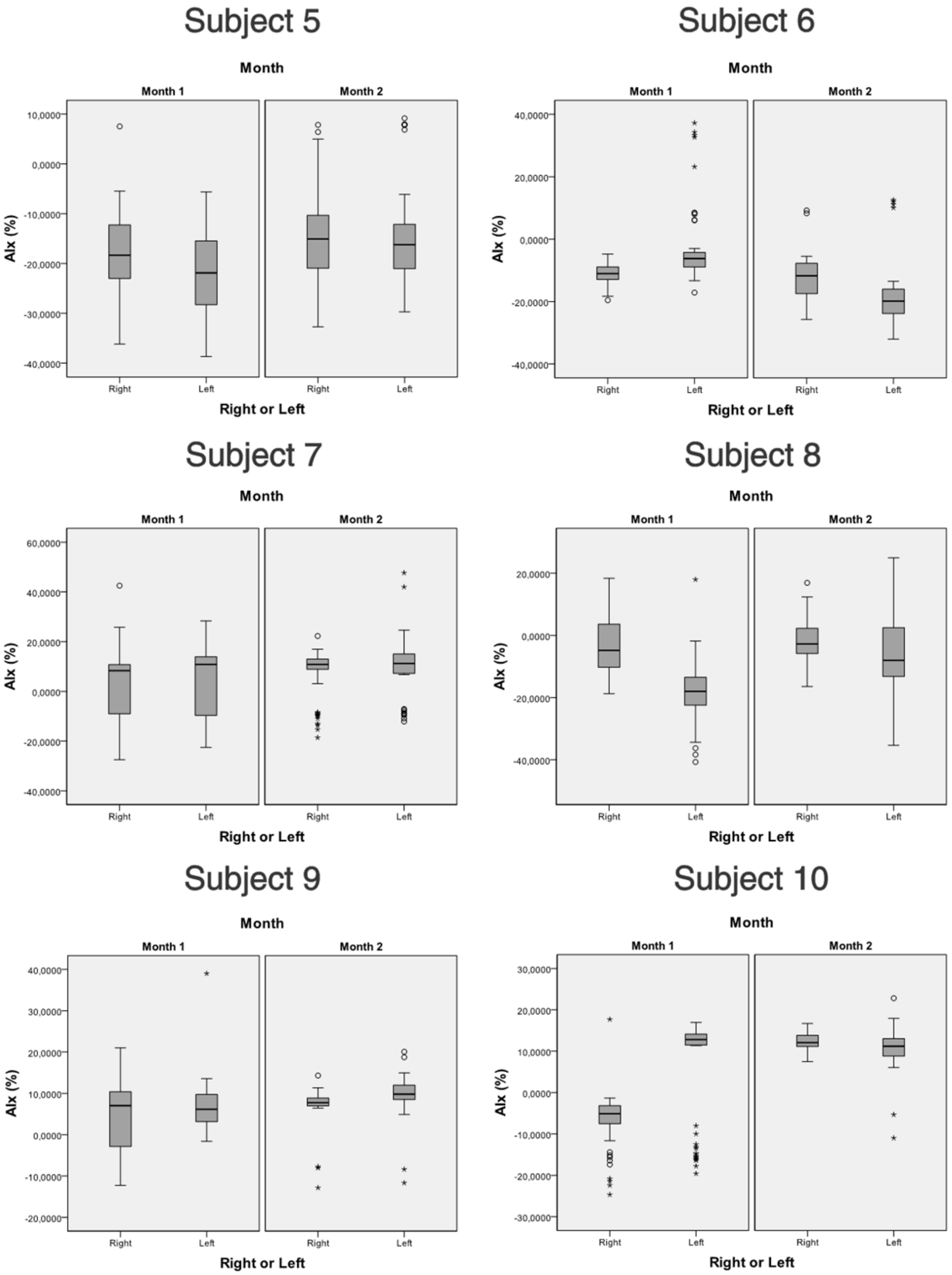


Figure A.2 – AIx boxplots of subjects 5-10 divided by month and carotid site;

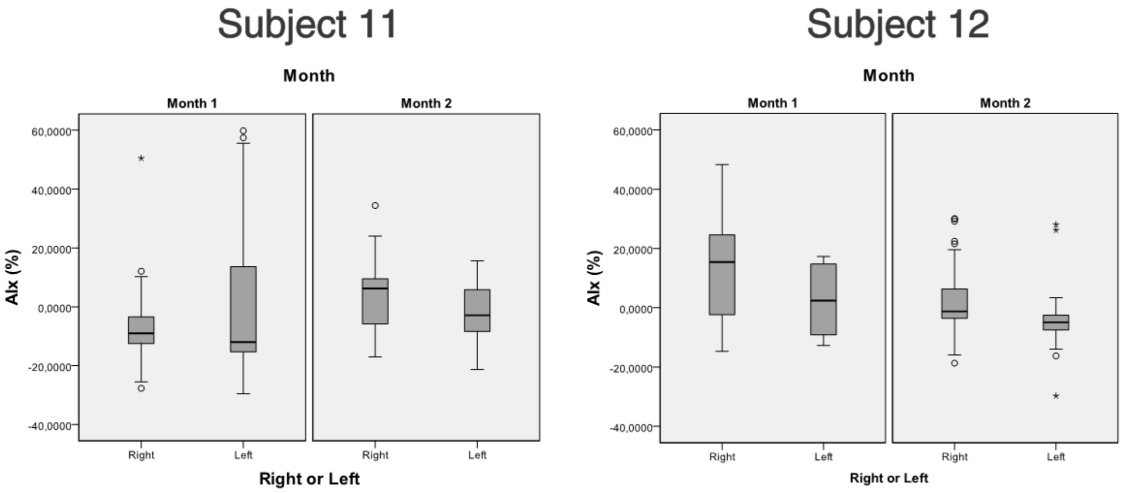


Figure A.3 – AIx boxplots of subjects 11 and 12 divided by month and carotid site;

Appendix B – Dataset I (2)

AIx table of each subject categorized by month, carotid site and trial

Table B.1 – AIx table of subjects 1-3 divided by month, carotid site and trial. Mean rank, Chi - square and P represent the conclusions given by the Kruskal Wallis test. A P – value of < .05 was considered as significant;

Subject	*Month*	*Carotid*	*Trial*	*Mean rank*	*Chi - square*	*P*
1	***1***	***Right***	**Trial 1**	23.00	**0.491**	**0.78**
			Trial 2	25.53		
			Trial 3	26.33		
		Left	**Trial 1**	17.23	**0.539**	**0.76**
			Trial 2	20.08		
			Trial 3	19.82		
	2	***Right***	**Trial 1**	39.10	**4.798**	**0.09**
			Trial 2	29.41		
			Trial 3	27.80		
		Left	**Trial 1**	19.00	7.323	< .05
			Trial 2	23.10		
			Trial 3	32.75		
2	***1***	***Right***	**Trial 1**	55.12	15.969	< .05
			Trial 2	42.86		
			Trial 3	28.93		
		Left	**Trial 1**	26.79	**0.952**	**0.62**
			Trial 2	32.07		
			Trial 3	28.00		
	2	***Right***	**Trial 1**	35.92	**2.606**	**0.27**
			Trial 2	27.35		
			Trial 3	35.00		
		Left	**Trial 1**	26.36	**3.915**	**0.14**
			Trial 2	24.87		
			Trial 3	35.25		
3	***1***	***Right***	**Trial 1**	71.97	56.407	< .05
			Trial 2	22.29		
			Trial 3	34.41		
		Left	**Trial 1**	14.25	10.106	< .05
			Trial 2	28.17		
			Trial 3	19.00		
	2	***Right***	**Trial 1**	36.23	**2.815**	**0.24**
			Trial 2	27.15		
			Trial 3	31.24		
		Left	**Trial 1**	24.90	7.886	< .05
			Trial 2	27.06		
			Trial 3	39.05		

Table B.2 – AIx table of subjects 4-6 divided by month, carotid site and trial. Mean rank, Chi - square and P represent the conclusions given by the Kruskal Wallis test. A P – value of < .05 was considered as significant;

Subject	*Month*	*Carotid*	*Trial*	*Mean rank*	*Chi - square*	*P*
4	***1***	***Right***	**Trial 1**	9.91	13.223	< .05
			Trial 2	13.40		
			Trial 3	24.00		
		Left	**Trial 1**	30.00	**5.664**	**0.06**
			Trial 2	17.50		
			Trial 3	31.71		
	2	***Right***	**Trial 1**	22.50	**0.201**	**0.90**
			Trial 2	22.87		
			Trial 3	21.00		
		Left	**Trial 1**	17.62	**1.370**	**0.50**
			Trial 2	13.50		
			Trial 3	14.43		
5	***1***	***Right***	**Trial 1**	76.27	**4.426**	**0.10**
			Trial 2	74.55		
			Trial 3	59.72		
		Left	**Trial 1**	32.34	18.710	< .05
			Trial 2	61.21		
			Trial 3	53.39		
	2	***Right***	**Trial 1**	52.86	7.272	< .05
			Trial 2	36.31		
			Trial 3	49.78		
		Left	**Trial 1**	26.97	24.987	< .05
			Trial 2	55.58		
			Trial 3	55.89		
6	***1***	***Right***	**Trial 1**	37.59	**3.322**	**0.19**
			Trial 2	27.84		
			Trial 3	30.94		
		Left	**Trial 1**	34.44	**0.079**	**0.96**
			Trial 2	34.48		
			Trial 3	33.05		
	2	***Right***	**Trial 1**	34.10	**4.503**	**0.10**
			Trial 2	23.52		
			Trial 3	28.69		
		Left	**Trial 1**	18.67	6.981	< .05
			Trial 2	26.24		
			Trial 3	33.00		

Table B.3 – AIx table of subjects 7-9 divided by month, carotid site and trial. Mean rank, Chi - square and P represent the conclusions given by the Kruskal Wallis test. A P – value of < .05 was considered as significant;

Subject	*Month*	*Carotid*	*Trial*	*Mean rank*	*Chi - square*	*P*
7	***1***	***Right***	**Trial 1**	75.73	18.806	< .05
			Trial 2	62.94		
			Trial 3	40.40		
		Left	**Trial 1**	56.46	6.661	< .05
			Trial 2	42.17		
			Trial 3	38.91		
	2	***Right***	**Trial 1**	36.23	**4.243**	**0.12**
			Trial 2	46.07		
			Trial 3	34.36		
		Left	**Trial 1**	28.00	**2.894**	**0.24**
			Trial 2	26.58		
			Trial 3	35.25		
8	***1***	***Right***	**Trial 1**	41.11	**3.134**	**0.21**
			Trial 2	46.83		
			Trial 3	36.08		
		Left	**Trial 1**	44.26	**1.087**	**0.58**
			Trial 2	50.00		
			Trial 3	51.10		
	2	***Right***	**Trial 1**	28.79	*6.138*	*0.046*
			Trial 2	30.33		
			Trial 3	18.07		
		Left	**Trial 1**	33.08	**1.175**	**0.56**
			Trial 2	37.55		
			Trial 3	31.33		
9	***1***	***Right***	**Trial 1**	23.73	10.339	< .05
			Trial 2	23.64		
			Trial 3	11.46		
		Left	**Trial 1**	10.63	0.542	**0.76**
			Trial 2	9.17		
			Trial 3	11.67		
	2	***Right***	**Trial 1**	11.33	8.045	< .05
			Trial 2	9.00		
			Trial 3	18.36		
		Left	**Trial 1**	25.25	4.687	**0.10**
			Trial 2	38.06		
			Trial 3	33.23		

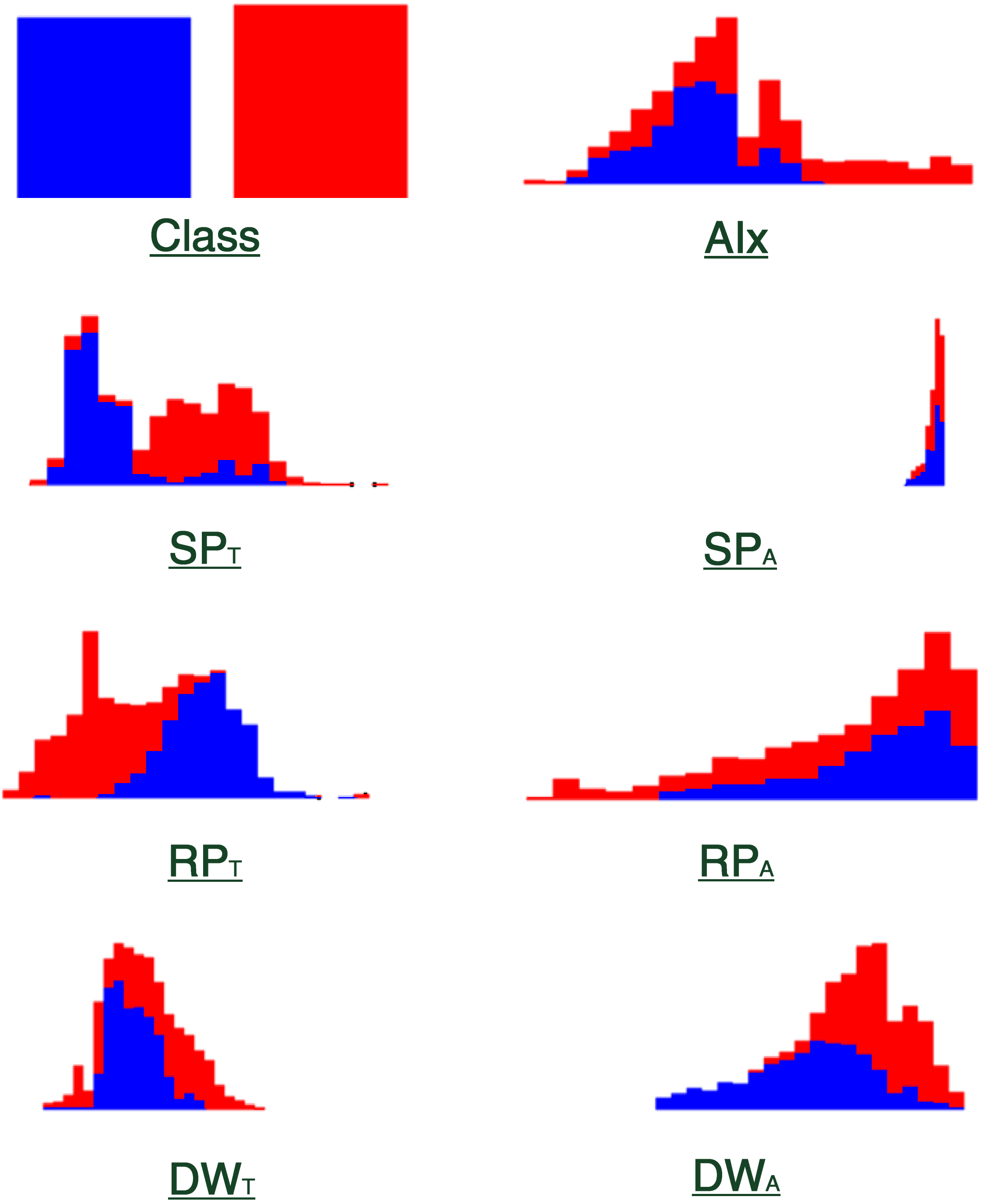
Table B.4 – AIx table of subjects 10-12 divided by month, carotid site and trial. Mean rank, Chi - square and P represent the conclusions given by the Kruskal Wallis test. A P – value of < .05 was considered as significant;

Subject	*Month*	*Carotid*	*Trial*	*Mean rank*	*Chi - square*	*P*
10	*1*	***Right***	**Trial 1**	18.47	7.754	< .05
			Trial 2	32.18		
			Trial 3	32.15		
		Left	**Trial 1**	25.96	12.369	< .05
			Trial 2	33.35		
			Trial 3	45.96		
	2	***Right***	**Trial 1**	35.92	**0.875**	**0.64**
			Trial 2	30.42		
			Trial 3	34.19		
		Left	**Trial 1**	28.81	**5.900**	**0.052**
			Trial 2	27.83		
			Trial 3	39.96		
11	*1*	***Right***	**Trial 1**	61.39	25.066	< .05
			Trial 2	23.43		
			Trial 3	54.00		
		Left	**Trial 1**	8.83	**0.626**	**0.73**
			Trial 2	8.00		
			Trial 3	10.40		
	2	***Right***	**Trial 1**	52.35	21.330	< .05
			Trial 2	20.00		
			Trial 3	38.68		
		Left	**Trial 1**	33.38	**1.264**	**0.53**
			Trial 2	28.09		
			Trial 3	28.20		
12	*1*	***Right***	**Trial 1**	20.33	**5.638**	**0.06**
			Trial 2	32.72		
			Trial 3	31.18		
		Left	**Trial 1**	5.00	**0.356**	**0.84**
			Trial 2	5.67		
			Trial 3	4.33		
	2	***Right***	**Trial 1**	23.10	**2.023**	**0.36**
			Trial 2	22.91		
			Trial 3	17.74		
		Left	**Trial 1**	45.62	**5.172**	**0.08**
			Trial 2	31.28		
			Trial 3	34.71		

Appendix C – Dataset III

Dataset III characterization

(note: all attributes except the class label are normalized between 0 and 1)



Class

AIx

SP$_T$

SP$_A$

RP$_T$

RP$_A$

DW$_T$

DW$_A$

Figure C.1 – Dataset III histograms of the following variables: Class, AIx, SP$_T$, SP$_A$, RP$_T$, RP$_A$, DW$_T$, DW$_A$;

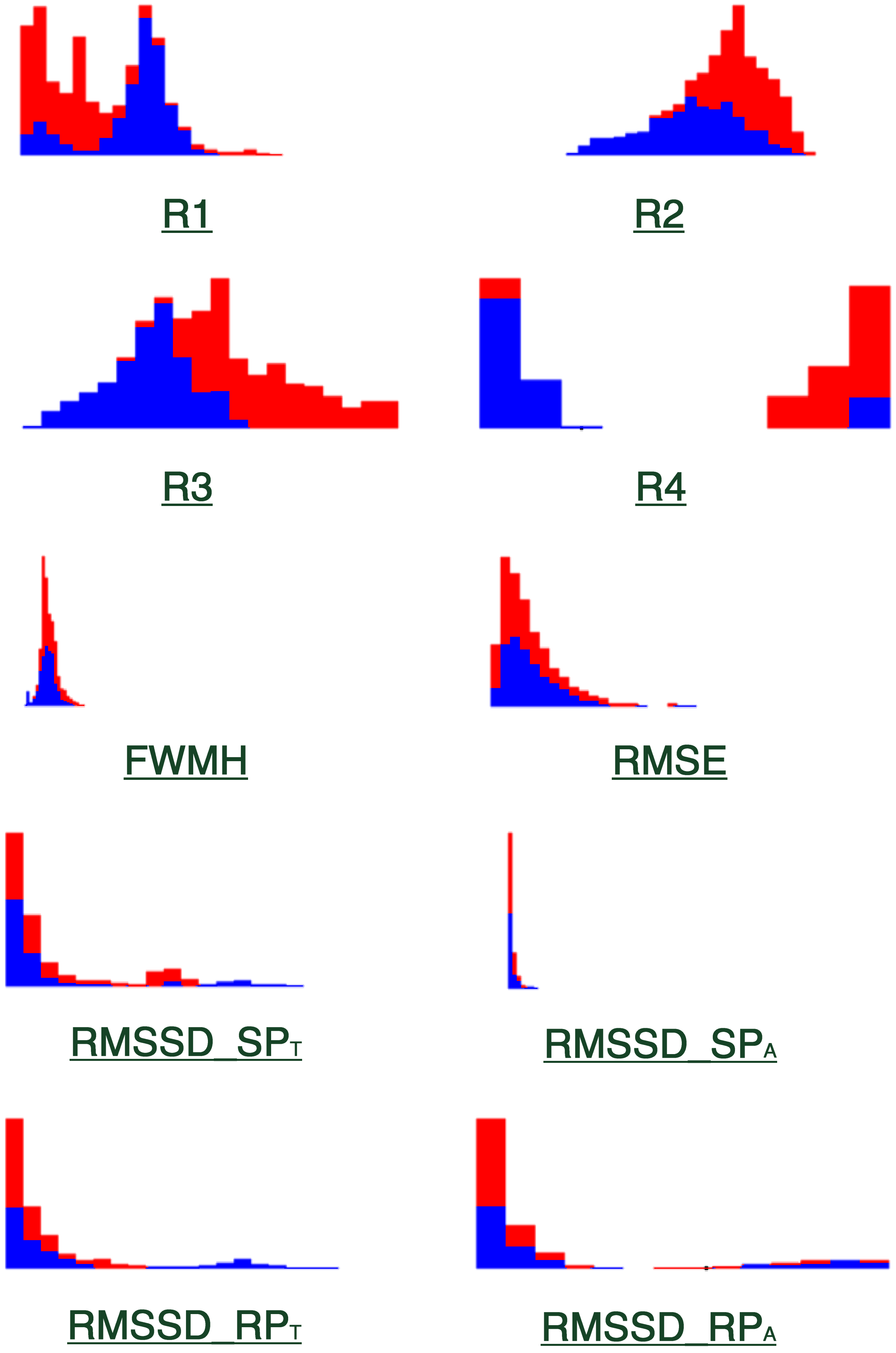


Figure C.2 – Dataset III histograms of the following variables: R1, R2, R3, R4, FWMH, RMSE, RMSSD_SP$_T$, RMSSD_SP$_A$, RMSSD_RP$_T$, RMSSD_RP$_A$;

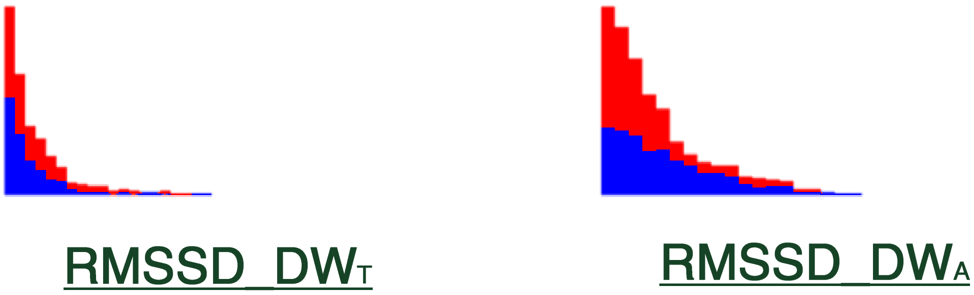


Figure C.3 – Dataset III histograms of the following variables: RMSSD_DW$_T$, RMSSD_DW$_A$;

Appendix D – Dataset IV

Dataset IV characterization

(Values expressed as mean ± SD)

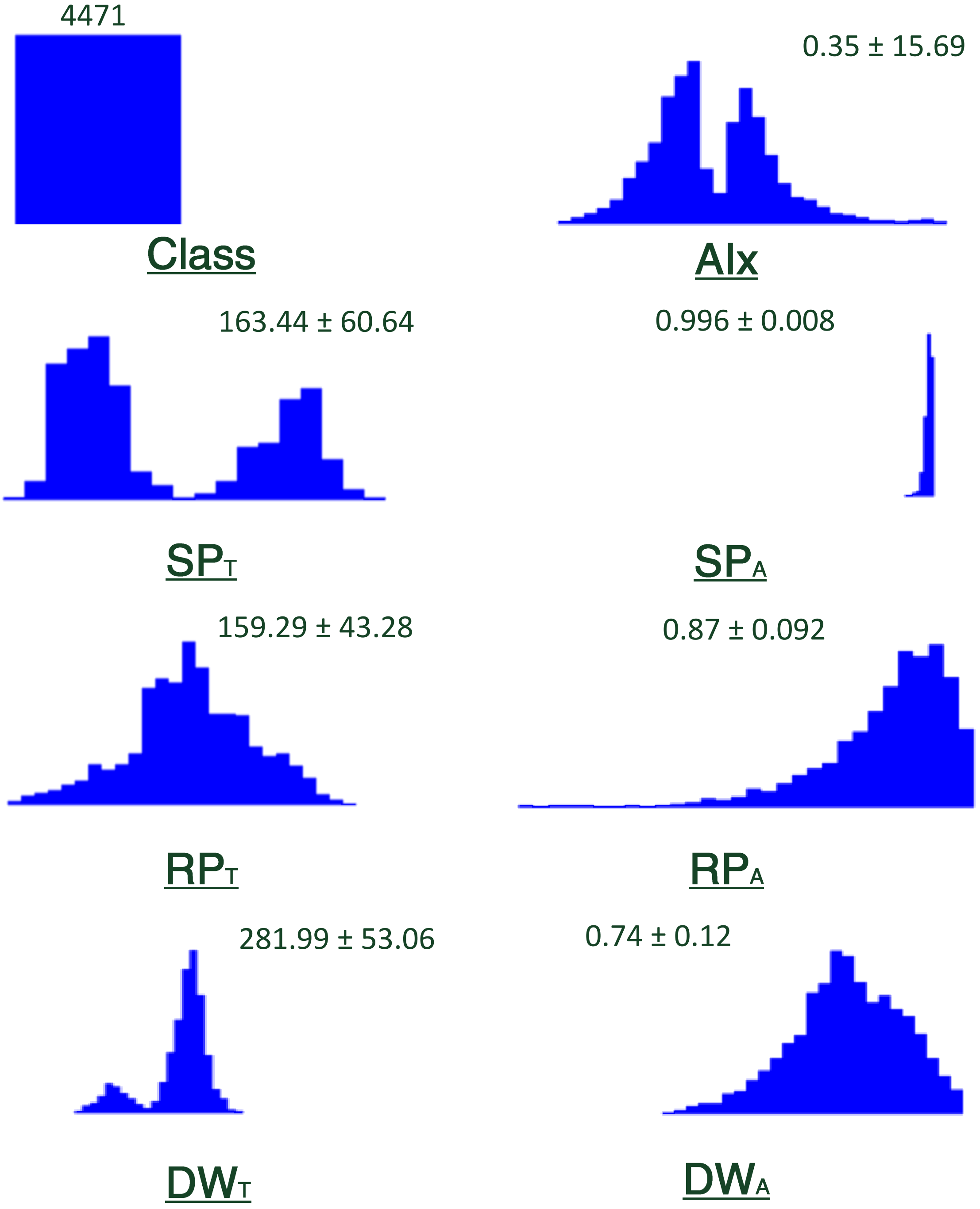


Figure D.1 – Dataset IV histograms of the following variables: Class, AIx, SP_T, SP_A, RP_T, RP_A, DW_T, DW_A;

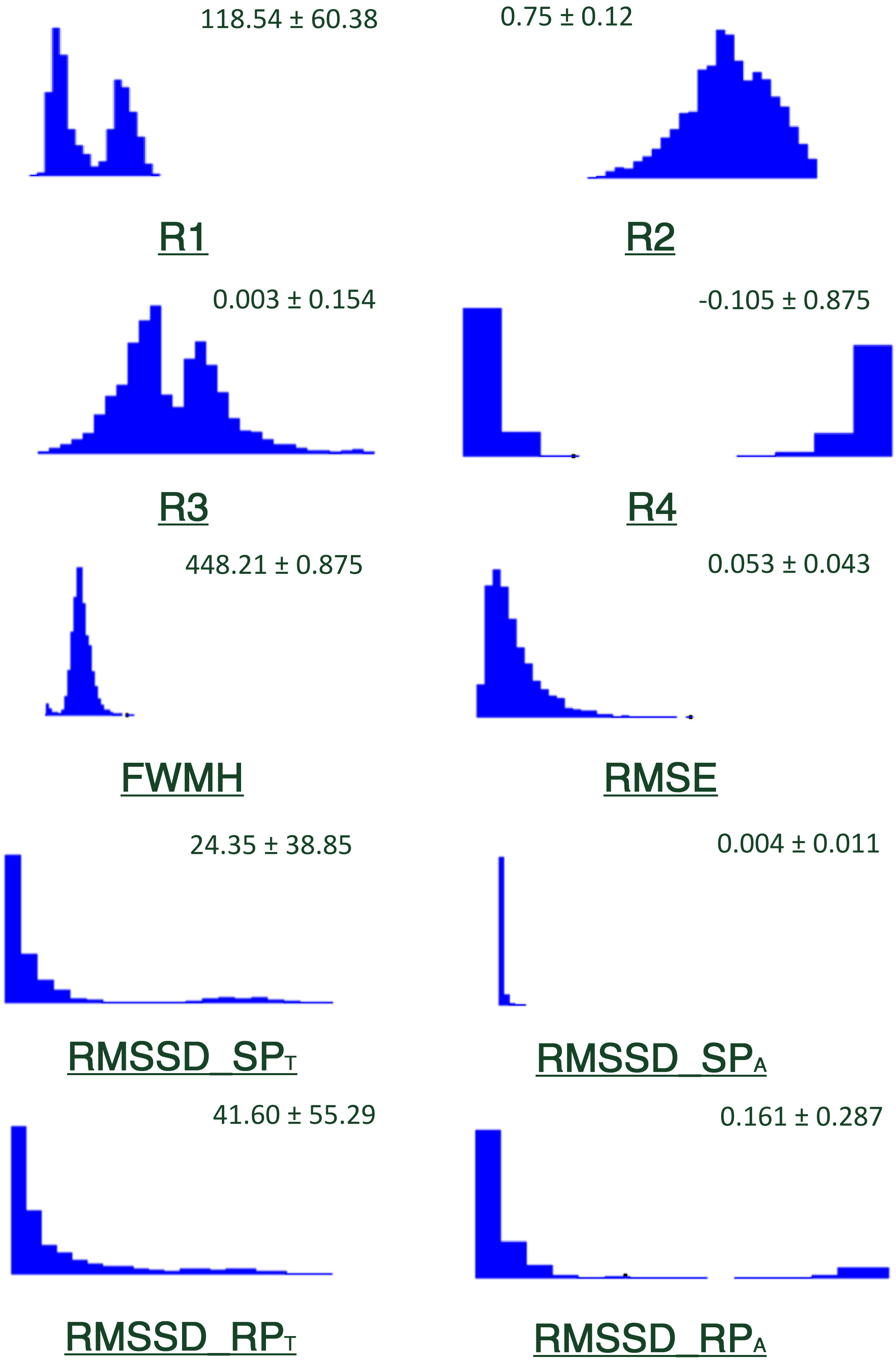


Figure D.2 – Dataset IV histograms of the following variables: R1, R2, R3, R4, FWMH, RMSE, RMSSD_SP$_T$, RMSSD_SP$_A$, RMSSD_RP$_T$, RMSSD_RP$_A$;

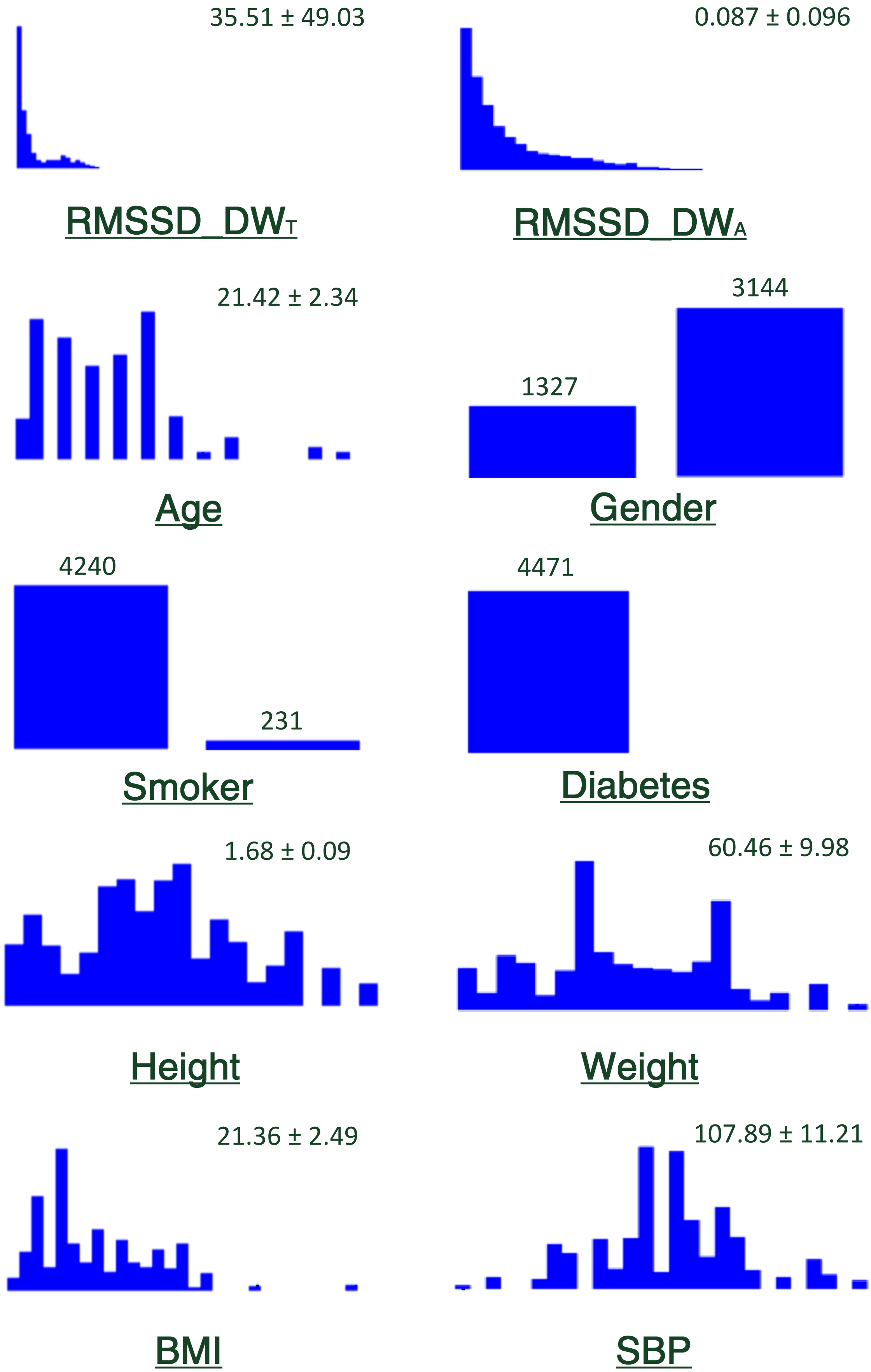


Figure D.3 – Dataset IV histograms of the following variables: RMSSD_DW$_{T}$, RMSSD_DW$_{A}$, Age, Gender, Smoker, Diabetes, Height, Weight, BMI, SBP;

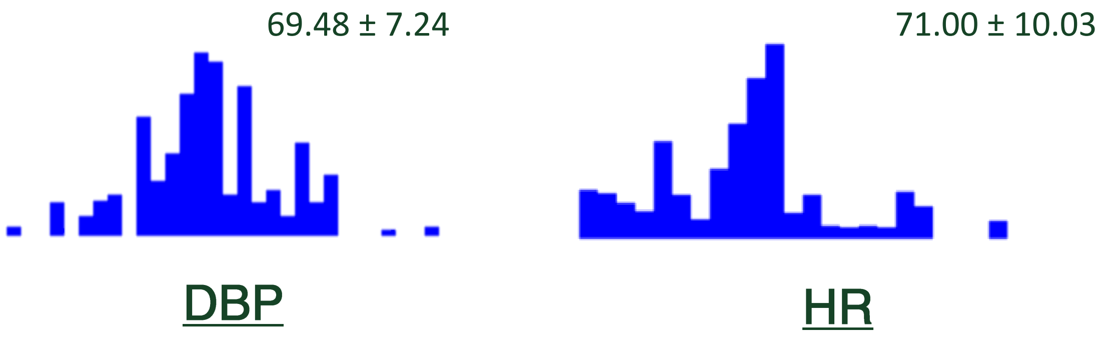


Figure D.4 – Dataset IV histograms of the following variables: DBP, HR;

Appendix E - Original Paper

I. Validation of a waveform delineator device for cardiac studies: repeatability and data mining analysis

Validation of a waveform delineator device for cardiac studies: repeatability and data mining analysis

V. G. Almeida[1], J. Borba[1], T. Pereira[1], H.C. Pereira[1,2], J. M. R. Cardoso[1], C Correia[1]
[1]Instrumentation Centre, Physics Department, University of Coimbra, Coimbra, Portugal
[2]ISA- *Intelligent Sensing Anywhere*, Coimbra, Portugal
vaniagalmeida@lei.fis.uc.pt

Abstract—**This paper envisages showing the potential of innovative non-invasive techniques based on affordable and easily operated instrumentation as well as user-friendly computer aided algorithms in the screening of cardiovascular (CV) diseases. These techniques are based on the assumption that arterial stiffness is currently an important predicator of the CV diseases development and can be assessed by analyzing the arterial pressure waveform (APW). A previously developed PZ based device for non-invasive APW capture is currently under test in clinical environment, using a heterogeneous population constituted by healthy and unhealthy subjects. A dedicated Matlab analysis tool was designed and developed to extract relevant information and further APW analysis. Several recordings of the APW in the same day and in consecutive months are being performed by trained observers, to evaluate its reproducibility. Data mining analysis is subsequently the last task where the Weka 3-6-5 package software is used. The usefulness of developing data mining algorithms for cardiovascular applications can benefit the CV screenings contributing for the early identification of arterial stiffness related patterns.**

Index Terms—Cardiovascular diseases, arterial stiffness, reproducibility tests, arterial pulse waveform, data mining analysis.

INTRODUCTION

Cardiovascular (CV) diseases are the number one cause of death globally, representing 30% of all global deaths. Their prevalence occur essentially in low and middle income countries due to higher exposition to risk factors and less prevention efforts than in high-income countries [1]. The importance of identifying the most important risk factor associated to the CV morbid events and develop efficient diagnostic tools to be used in early stages of development is evident.

Arterial stiffness has been associated with CV diseases development by several authors. Its occurrence can denote alterations in the mechanical properties of the arteries, generally related with the decay of elasticity in the arterial wall fibers [2, 3]. Much interest has been paid to the arterial pressure waveform (APW) analysis over the last years, using methods that accurately extract important clinical information [4]. Many parameters have already been proposed in the literature, such as the Augmentation Index (AIx). This parameter is described as the augmentation of systolic pressure peak imparted to the APW by the propagating reflected wave [4].

The use of piezoelectric sensors in APW measurements has been reported by several authors. In a previously work, a non-invasive device for APW monitoring was developed and tested in laboratory and *in vivo* data applications [5, 6], with good accuracy results in the signal reproduction.

The recent emergence of computer-aided diagnosis (CAD) technologies has claimed for work on innovative algorithms to assist health professionals in interpreting and in building new insights from cardiac data. CAD algorithms have been already proposed in the literature for coronary arterial disease detection [7, 8], or electrocardiogram (ECG) abnormalities [9]. The determination of the key indicator parameters from the APW signals [10] can become an important tool in CV screening trials, contributing for the early identification of arterial stiffness related patterns.

GOALS

This projects aims at the development of efficient diagnostic tools based on non-invasive devices that can be easily operated in diagnostic trials. In previous tasks, important hardware and firmware developments contributed to the development of a non-invasive PZ probe. The use of this probe can be a convenient and affordable solution to assess the hemodynamic condition.

Currently, we are working on the clinical validation of the PZ probe, which includes repeatability and reproducibility tests in a controlled medical environment. Likewise, data mining techniques are used in the conviction that a full APW analysis can contribute to arterial stiffness pattern recognition, which could prove as an outstanding achievement in entrusting a premature and correct diagnosis of CV diseases.

TEAM AND INSTITUTIONS

The Electronics and Instrumentation Group (GEI) is located on the Physics Department is strongly involved in the development of instrumentation and processing techniques in the fields of Biomedical Engineering. In the last years, an interdisciplinary expertise in biomedical instrumentation required to meeting the challenges outlined above was developed. The strength of this project relies on the simultaneous integration of instrumentation previously developed and innovative algorithms to signal analysis.

The Coimbra University Hospital and Coimbra's Hospital Center are important partners where clinical tests are being performed.

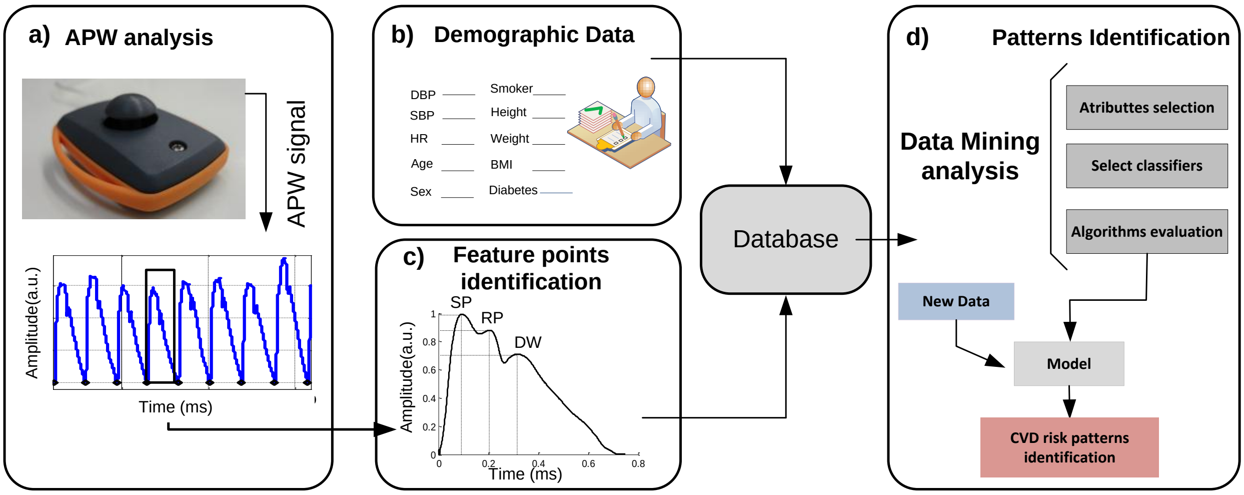


Fig.1 Schematic representation of the main tasks in this project: a) APW capture, b) demographic data registration, c) prominent points analysis, d) algorithms development for patterns identification.

Implementation

The main tasks performed in this project are schematized in Figure 1. The non-invasive probe used in clinical validation is shown in Figure 1(a) (upper row). A segment data composed by 8 pulses is also shown after baseline removal (lower row). Each APW is then analyzed and morphological parameters are obtained from it. This information is then stored in a database alongside with the subject's demographic data (Figure 1(b) and (c)). Finally, previously selected classification algorithms are tested and evaluated in order to be used in the model prediction development (figure 1 (d)), where the new data from undiagnosed subjects are studied.

A. *Subjects and study protocol*

In order to prove the importance of the APW analysis we are studying two groups. Group I is used for repeatability and reproducibility measurements and Group II for data mining analysis. The demographic data of both groups are presented in Table I.

APWs were recorded at the sampling rate of 1kHz using the non-invasive PZ probe previously developed [5]. The probe, shown on Figure 1c), is placed over the carotid artery, and is held by a collar to avoid noise and artifacts that arise from the interaction between the probe and the human operator's hand.

Age, sex, weight, height, smoking habits and diabetes history were registered and BMI was later calculated for all subjects. Blood pressure (BP) and heart rate (HR) values were measured in the left arm with an automated digital oscillometric sphygmomanometer (Omron M6 Comfort). The study was approved by the Committees of the Coimbra University Hospital and Coimbra's Hospital Center where the data acquisition were taken.

Dataset I

This group is composed by 8 healthy subjects (5 female 3 male). All of the subjects were non – smokers, and had no documented history of CV disorders or diabetes. All measurements were made in a similar time of the day on each month, at the same temperature controlled room (22 – 23ºC). Subjects remained quiet and seated on a comfortable chair during the measurements.

In two successive months, at least four measurements separated by 1 – minute intervals were made for each subject, by one observer. As a quality control procedure, the four best trials (two per month) were chosen for statistical analysis. Then, 20 continuous pulses were randomly chosen from each trial summing 80 pulses for each subject. and 740 pulses in the whole dataset.

TABLE I. Demographic data

Variable	*Dataset I*	*Dataset II*	
	Mean ± SD	*Mean ± SD*	
		Sub-group I	*Sub-group II*
Age (years)	23,88 ± 2,85	24,16 ± 3,86	58,16 ± 11,77
Sex (male/female)	3/5	12/13	12/13
Smoker (yes/no)	0/8	3/22	4/21
Weight (kg)	59,13 ± 8,69	65,28 ± 10,42	75,25 ± 10,42
Height (m)	1,66 ± 0,07	1,70 ± 0,06	1,64 ± 0,08
BMI (kg/m^2)	21,23 ± 1,91	22,48 ± 2,69	28,18 ± 5,24
SBP (mmHg)	102,63 ± 8,4	110,20 ± 11,94	161,05 ± 17,88
DBP (mmHg)	69,13 ± 10,60	71,00 ± 11,17	94,90 ± 11,73
HR (beats/min)	70,63 ± 14,84	68,44 ± 10,72	68,73 ± 6,48

Dataset II

This group is composed by a heterogeneous population (N=50), divided in two sub-groups. Sub-group I is constituted by 25 healthy volunteers without documented history of CV disorders. Sub-group II is constituted by 25 hypertensive subjects. The hypertensive data were acquired during hospitalization but prior to taking any medication.

B. *Feature analysis*

A set of morphological features were chosen to be used in APW characterization [11]. The list includes time and amplitude position of the most important prominent points: systolic peak (SP), dicrotic wave (DW) and reflection point (RP), represented in Figure 1c).The relevance of SP amplitude analysis is negligible due to the previously normalization.

The time and amplitude analysis is essential in the study of the most important predictive patterns that are used in data mining analysis. The elimination of low predictive parameters as DW was possible due to the similar values in both groups.

C. *Statistical analysis*

As a "*a priori*" statistical analysis, all data were tested for normality, applying the Kolmogorov – Smirnov one-sample test, where the maximum difference between the sample cumulative distribution and the hypothesized cumulative distribution are compared. All continuous variables in Table II presented a non – normal distribution.

For each parameter, the mean of each trial was computed and compared with the following month using the Bland Altman method [12] to assess the repeatability and reproducibility of the PZ sensor. This method calculates the mean difference between measurements (the 'bias') and 95% limits of agreement as the mean difference (1,96 SD). It is expected that the 95% limits include ≈ 95% of the differences.

All data were analyzed as Mean ± SD with Predictive Analytics Software Statistics 18 (SPSS, Inc, Chicago, IL). The level of statistical significance was set at $p < 0.05$ for all analyses.

D. *Data mining analysis*

The Weka system was selected as the analysis tool due to its efficiency, versatility and affordability. A set of Weka classifiers (Random Forest, J48, and JRIP) were chosen and applied in our dataset (hypertensive and normal subjects) with objective of choose the most accurate. The number of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) were determined for each classifier.

ROC (receiver operating characteristic) curves were plotted for all classifiers (sensitivity vs. 1-specificity). A large area under the ROC curve (AUC=1) reflects superior classifier discrimination between the patterns. The accuracy (ACC) of the system is determined by equation (1), where C= TP + TN + FP + FN.

$$ACC = \frac{TP + TN}{C} \tag{1}$$

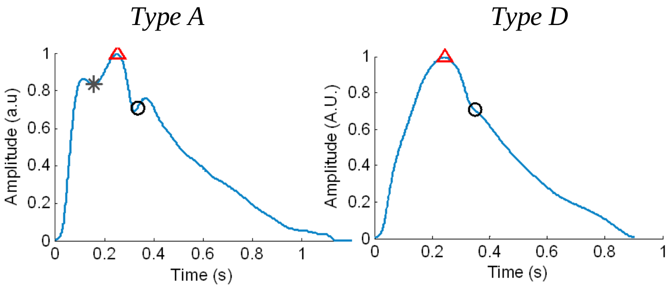


Fig.2 Two waveform types: type A and D, taken from the dataset II.

Results from ACC and AUC are used to choose the best model predictors. This model is used in the patterns identification from undiagnosed subjects.

Results

E. *In vivo APW*

The *in vivo* APW analysis includes different APW waveform types, according classification purposed by Murgo *et. al.* [13]. The Figure 2 illustrates two waveform types, which were taken from our dataset, from a middle age and an old subject, respectively

F. *Statistical Analysis*

Descriptive data analysis for both datasets is shown of Table II. Bland – Altman plots were constructed using dataset I shown in Figure 3 for time and amplitude measurements, as well as AIx.

Repeatability tests

Even if only Figure 3(d) includes > 95% of the differences, all figures show a reasonable value of agreement of 93.75% for Figure (b), (c), (e) and (f) and 87,5% for Figure (a). This suggests that even though further tests have to be done to reach sufficient agreement, a good repeatability level has already been reached. It is worth of notice the gap between values in Figure 3(b). This represents the separation between type the APWs type A and type C [13].

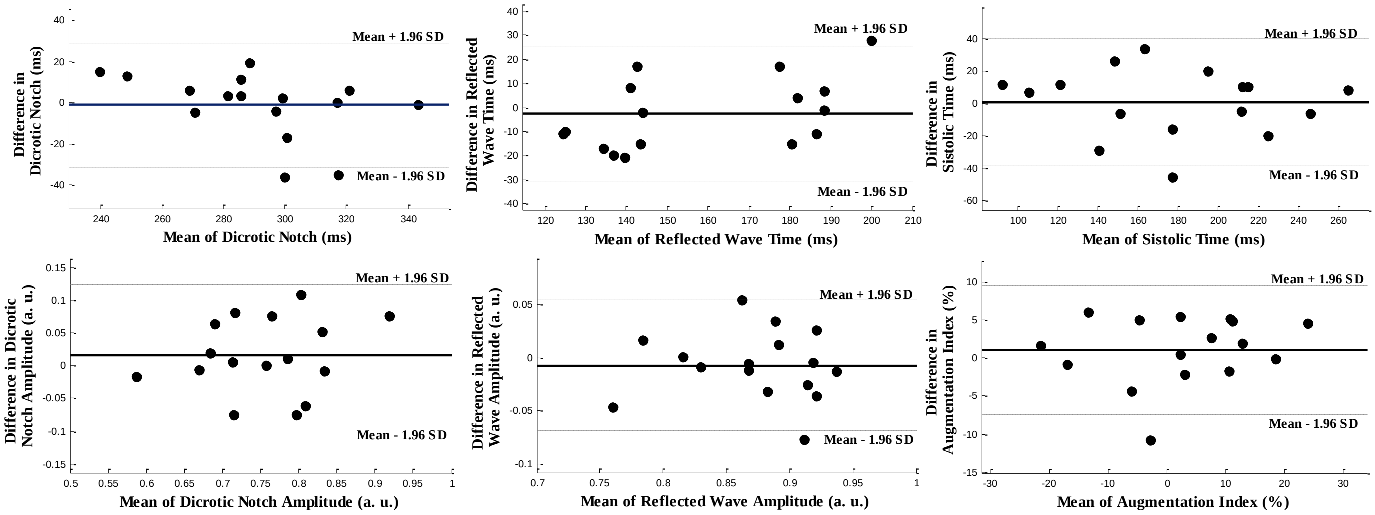


Fig.3 Bland Altman plots from reproducibility tests.

Features analysis

Table II shows the time and amplitude related parameters obtained for each group and sub-group, the values between group I and sub-group I are similar, as expected, due to the same healthy subject constitution in both groups. SP occurs later in hypertensive subjects (sub-group II) contrasting with early occurring of RP in these subjects. DW arrival time is similar for both groups, as referred below.

TABLE II. FEATURE POINT VALUES OBTAINED FOR ALL GROUPS.

Variable	***Group I (Mean ± SD)***	***Group II (Mean ± SD)***	
		Sub-group I	***Sub-group II***
SP time (ms)	177,68 ± 62,06	159,09 ± 57,69	222,61 ± 53,83
DN time (ms)	291,47 ± 39,83	294,48 ± 33,63	300,28 ± 50,99
RP time (ms)	158,38 ± 36,40	188,12 ± 39,87	119,37 ± 39,23
DN amplitude (ratio)	0,75 ± 0,11	0,70 ± 0,12	0,84 ± 0,09
RP amplitude (ratio)	0,87 ± 0,08	0,86 ± 0,10	0,74 ± 0,17

G. Data mining analysis

The AUC values obtained for each one of the three classifiers were: 0.994 for Random Forest, 0.961 for J48 and 0.965 for JRIP classifier. The results demonstrated the high accuracy values obtained for all of the classifiers, in special the performance of Random Forest classifier.

The ACC results, shown in Table II, confirm the superior performance of Random Forest classifier. The J48 and JRIP have similar performance values.

TABLE II. ACUURACY RESULTS FOR THE TESTED ALGORITHMS.

Classifier	RANDOM FOREST	J48	*JRIP*
ACC (%)	*96.95*	*95.90*	*94.78*
AUC	*0.994*	*0.961*	*0.965*

CONCLUSIONS

The nature of arterial wave propagation of incident and reflected waves plays a major role in the determination of important parameter indicators, which serve as health status predictors of the CV system. The APW signal is an interesting signal to this purpose and can be easily and affordable obtained by the non invasive instrumentation developed in our group. The reliability and repeatability tests demonstrated its clinical value.

The use of the data mining tools in the biomedicine should bring revolutionary impact to this field. The study of biomedical processes is heavily based on the identification of understandable patterns which are present in the data. These patterns may be used for diagnostic or prognostic purpose. The performance values obtained by the model predictors allow anticipating the good results in its application.

PLANNED DEVELOPMENTS

Currently, we are running a set of clinical tests to improve the number of dataset subjects, to prove that this system is a valid and low-cost alternative to the standard devices. Parameters from other cardiac setups, as biochemical analysis, will be included in attributes list in future trials, to address a more accurate diagnosis tool for CV risk assessment.

Also, the use of neural networks as superior data mining algorithms has been reported by some authors, with good results so far, and there are plans for implementing different types of neural networks as data mining classifiers in this project.

ACKNOWLEDGMENT

Authors acknowledge *Fundação para a Ciência e Tecnologia* for funding SFRH/BD/61356/2009 and PTDC/SAU-BEB/100650/2008, project developed under the initiative of QREN, funding by UE/FEDER, through COMPETE *-Programa Operacional Factores de Competitividade*. The authors also thanks to: ISA-Intelligent Sensing Anywhere, Coimbra University Hospital and Coimbra's Hospital Center.

REFERENCES

(2011) World health organization website. [Online]. Available: http://www.who.int/cardiovascular_diseases/en/

Laurent, S., J. Cockcroft et al. "Expert consensus document on arterial stiffness: methodological issues and clinical applications." European Heart Journal 27: 2588 – 2605

Cheung, Y.-F (2010). "Arterial Stiffness in the Young: Assessment, Determinants and Implications." The Korean Society of Cardiology Journal **40**(4): 153 – 162

Avolio, A. P., M. Butlin, et al. (2010). "Arterial blood pressure measurement and pulse wave analysis--their role in enhancing cardiovascular assessment." Physiol Meas 31(1): R1-4

Almeida, V., T.Pereira, et al. (2011). "Piezoelectric probe for pressure waveform estimation in flexible tubes and its application to the cardiovascular system" Sensors and Actuators A 169: 217 - 226

Almeida, V., T.Pereira, et al. (2008). "A real time cardiac monitoring system – Arterial pressure waveform capture and analysis", unpublished.

M. Tsipouras, et. al., Automatic Creation of Decision Support Systems: Application and Results in the Cardiovascular Diseases Domain, The Journal on Information Technology in Healthcare 2006; 4(4): 222–230

M. G. Tsipouras, Automated Diagnosis of Coronary Artery Disease Based on Data Mining and Fuzzy Modeling, IEEE Transactions On Information Technology In Biomedicine, VOL. 12, NO. 4, JULY 2008

Q. Y. Lee, et. al. Multivariate classification of systemic vascular resistance using photoplethysmography. Physiol. Meas. 32 (2011)1117-1132.

A multi-parametric arterial pressure waveform analysis based on data mining approaches (under submission).

V. G. Almeida, P. Santos, E. Figueiras, E. Borges, T. Pereira, J. Cardoso, C. Correia, C. Pereira, Hemodynamic features extraction from a new arterial pressure waveform probe. Proceedings of the 4th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2011), Rome, Italy, 26-29 January 2011

Bland, J. M. and D. G. Altman (1986). "Statistical Methods for Assessing Agreement Between Two Methods Of Clinical Measurement." Lancet 1: 307 - 310

Murgo, J.P.,Westerhof, N.,Giolma J.P., Altobelli, S.A. (1980), Aortic input impedance in normal man: relationship to pressure wave forms. Circulation, 62, 105-116

References

[1] – Mendis, S., Puska, P., Norrving, B., Global Atlas on cardiovascular disease prevention and control, World Health Organization, 2011;

[2] – Laurent, Stéphane, *et al.*, Expert consensus document on arterial stiffness: methodological issues and clinical applications, *European Heart Journal* 27:2588-2605, 2006;

[3] – Fernhall, B., Agiovlasitis, S., Arterial function in youth: window into cardiovascular risk, *J Appl Physiol*, 105:325-33, 2008;

[4] – Cheung, Yiu – Fui, Arterial Stiffness in the Young: Assessment, Determinants and Implications, *Korean Circ J*, 40(4):153-162, 2010;

[5] – Lee, HY, Oh, BH, Aging and Arterial Stiffness, *Circ J.*, 74(11):2257-2262, 2010;

[6] – Laurent, S., Boutouyrie, P., Asmar, R., Gautier, I., Laloux, B., Guize, L., Ducimetiere, P., Benetos, A., Aortic stiffness is an independent predictor of all – cause and cardiovascular mortality in hypertensive patients, *Hypertension,* 37:1236-1241, 2001;

[7] – Blacher, J., Pannier, B., Guerin AP., Marchais, SJ., Safar, ME., London, GM., Carotid arterial stiffness as a predictor of cardiovascular and all-cause mortality in end – stage renal disease, *Hypertension* 32:570-574, 1998;

[8] – Boutouyrie, P., Tropeano, AI., Asmar, R., Gautier, I., Benetos, A., Lacolley, P., Laurent, S., Aortic stiffness is an independent predictor of primary coronary events in hypertensive patients: a longitudinal study, *Hypertension* 39:10-15, 2002;

[9] – Laurent, S., Katsahian, S., Fassot, C., Tropeano, AI., Gautier, I., Laloux, B., Boutouyrie, P., Aortic stiffness is an independent predictor of fatal stroke in essential hypertension, *Stroke,* 34:1203-1206, 2003;

[10] – McLaughlin, J., Mcneill, M., Braun, B. McCormack, PD., Piezoelectric sensor determination of arterial pulse wave velocity, *Physiological Measurement*, 24:693-702, 2003;

[11] – Pereira, HC., Pereira, T., Almeida, V., Borges, E., Figueiras, E., Simões, EJB., Malaquias, JL., Cardoso, JMR., Correia, CMB., Characterization of a double probe for local pulse wave velocity assessment, *Physiological Measurement,* 31:1449-1465, 2010;

[12] – Clemente, F., Arpaia, P., Cimmino, P., A piezo – film – based measurement system for global haemodynamic assessment, *Physiological Measurement,* 31:697-714, 2010;

[13] – Almeida, VG., Pereira, T., Borges, E., Cardoso, JMR., Correia, C., Pereira, HC., A Real Time Cardiac Monitoring System – Arterial pressure waveform capture and analysis, In *Proceedings of the 1st International Joint Conference on Pervasive and Embedded Computing and Communication Systems (PECCS 2011)*;

[14] – Almeida, VG., Pereira, H.C., Pereira, T., Figueiras, E., Borges, E., Cardoso, JMR., Correia, C., Piezoelectric probe for pressure waveform estimation in flexible tubes and its applications to the cardiovascular system, *Sensors and Actuators A: Physical*, 169:217-226;

[15] – Crilly, M., Coch, C., Bruce, M., Clark, H., Williams, D., Indices of cardiovascular function derived from peripheral pulse wave analysis using radial applanation tonometry: a measurement repeatability study, *Vascular Medicine*, 12:189-197, 2007;

[16] – Bland, JM., Altman, DG., Statistical methods for assessing agreement between two methods of clinical measurement, *The Lancet*, 327(8476):307-310, 1986;

[17] – Yoo, I., Alafaireet, P., Marinov, M., Pena-Hernandez, K., Gopidi, R., Chang, JF., Hua, L., Data Mining in Healthcare and Biomedicine: A Survey of the Literature, *J Med Syst*, 36:2431-2448, 2012;

[18] – Kumari, M., Godara, S., Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction, *IJCST*, 2(2):304-308;

[19] – Paredes, S., Rocha, T., de Carvalho, P., Henriques, J., Harris, M., Morais, J., Long term cardiovascular risk models combination, *Computer Methods and Programs in Biomedicine,* 101:231-242, 2011;

[20] – Shah, AS., Dolan, LM., Gao, Z., Kimball, TR., Urbina, EM., Clustering of Risk Factors: A Simple Method of Detecting Cardiovascular Disease in Youth, *Pediatrics*, 127(2):312-318, 2011;

[21] – Pereira, HC., Cardioaccelerometery, Universidade de Coimbra, Setembro 2007;

[22] - Almeida, VMG., Hemodynamic Parameters Assessment – An Improvement Of Methodologies, Universidade de Coimbra, Setembro 2009;

[23] – Pereira, T., Methodologies for Hemodynamic Parameters Assessment, Universidade de Coimbra, Setembro 2009;

[24] – Vieira, J., Algorithm development for physiological signals analysis and cardiovascular disease diagnosis – A data mining approach, Universidade de Coimbra, 2011;

[25] – Almeida, V., Santos, P., Figueiras, E., Borges, E., Pereira, T., Cardoso, J., Correia, CMBA., Pereira, H.C., Hemodynamic features extraction from a new arterial pressure waveform probe, *Biosignals*, Rome, 2011;

[26] – Almeida, VG., Vieira, J., Santos, P., Pereira, HC., Pereira, T., Borges, E., Cardoso, J., Correia, C., A multi – parametric arterial pressure waveform analysis based on data mining approaches, unpublished;

[27] – Shen, Z., Clarke, M., Jones, R., Alberti, T., A new neural network structure for detection of coronary heart disease, *Neural Computing & Applications,* 3(3):171-177, 1995;

[28] – Das, R., Turkoglu, I., Sengur, A., Effective diagnosis of heart disease through neural network ensembles, *Expert Systems with Applications,* 36:7675-7680, 2009;

[29] – Al – Shayea, QM., Artificial Neural Networks In Medical Diagnosis, *IJCSI*, 8(2):150-154;

[30] – Patil, SB., Kamaraswamy, YS., Intelligent and Effective Heart Attack Prediction System Using Data Mining and Artificial Neural Network, *European Journal of Scientific Research,* 31(4):642-656, 2009;

[31] – Raut, R., Dudul, SV., Intelligent Diagnosis of Heart Diseases using Neural Networks Approach, *International Journal of Computer Applications*, 1(2):97-102, 2010;

[32] – Gorunescu, F., Gorunescu, M., Saftoiu, A., Vilmann, P., Belciug, S., Competitive/collaborative neural computing system for medical diagnosis in pancreatic cancer detection, *Expert Systems*, 28(1):33-48, 2011;

[33] – Bronzino, JD., Biomedical Engineering Handbook, CRC Press LLC, 2000;

[34] – Guyton, AC., Hall, JH., Textbook of Medical Physiology (11th Edition), Saunders Co., 2005;

[35] – Stranding, S., Gray’s Anatomy – The Anatomical Basis of Clinical Practice, Elsevier, 2008;

[36] – Claridge, MWC., Clinical assessment of arterial stiffness, The University of Birmingham, 2009;

[37] – Marieb, EN., Hoehn, K., Human Anatomy & Physiology (7th Edition), Benjamin Cummings, 2007;

[38] – Oliver, JJ., Webb, DJ., Noninvasive assessment of arterial stiffness and risk of atherosclerotic events, *Arterioscler Thromb Vasc Biol*, 23(4):554-566, 2003;

[39] – Nichols, WW., Clinical Measurement of Arterial Stiffness Obtained from Noninvasive Pressure Waveforms, *AJH*, 18:3S-10S, 2005;

[40] – O'Rourke, MF., Staessen, JA., Vlachopoulos, C., Duprez, D., Plante, GE., Clinical Applications of Arterial Stiffness; Definitions and Reference Values, *AJH,* 15:426-444, 2002;

[41] – Murgo, JP., Westerhof, N., Giolma, JP., Altobelli, SA., Aortic input impedance in normal man: relationship to pressure wave forms, *Circulation,* 61:105-116, 1980;

[42] – Rönnback, M., Arterial stiffness and cardiovascular risk factors, University of Helsinki, 2007;

[43] – Korteweg, DJ., Uber die Fortpflanzungsgeschwindigkeit des Schalles in Elastichen Rohren, *Annalen der Physik.* 241(12):525-542, 1878;

[44] – Bramwell, JC., Hill, AV., The velocity of pulse wave in man, *Proceedings of the Royal Society of London, Series B,* 93(652):298-306, 1922;

[45] – Nurnberger, J., Keflioglu-Scheiber, A., Opazo Saez, AM., Wenzel, RR., Philipp, T., Schafers, RF., Augmentation index is associated with cardiovascular risk, *J Hypertens,* 20(12): 2407-2414, 2002;

[46] – Janner, JH., Godtfredsen, NS., Ladelund, S., Vestbo, J., Prescott, E., Aortic Augmentation Index: Reference Values in a Large Unselected Population by Means of the SphygmoCor Device, *American Journal of Hypertension,* 23(2):180-185, 2010;

[47] – Chung, JW., Lee, YS., Kim, JH., Seong, MJ., Kim, SY., Lee, JB., Ryu, JK., Choi, JY., Kim, KS., Chang, SG., Lee, GH., Kim, SH., Reference Values for the Augmentation Index and Pulse Pressure in Apparently Healthy Korean Subjects, *Korean Circ J,* 40(4):165-171, 2010;

[48] – Wimmer, NJ., Townsend, RR., Joffe, MM., Lash, JP., Go, AS., Correlation between pulse wave velocity and other measures of arterial stiffness in chronic kidney disease, *Clin Nephrol*, 68(3):133-143, 2007;

[49] – McEniery, CM., Yasmin., Hall, IR., Qasem, A., Wilkinson, IB., Cockcroft, JR., Normal Vascular Aging: Differential Effects on Wave Reflection and Aortic Pulse Wave Velocity: The Anglo-Cardiff Collaborative Trial (ACCT), *J Am Coll Cardiol,* 46:1753-1760, 2005;

[50] – Barrett, KM., Brott, TG., Carotid Artery Stenting Versus Carotid Endarterectomy: Current Status, *Neurol Clin,* 24:681-695, 2006;

[51] – Repeatability and Reproducibility, *Engineered Software Inc,* 1999;

[52] – Taylor, BN., Kuyatt, CE., Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, National Institute of Standards and Technology, United States of America, 1994;

[53] – Pallant, Julie, SPSS Survival Manual (Version 12), *Allen&Unwin,* 2005;

[54] – Walpole, RE., Myers, RH., Myers, SL., Ye, K., Probability and Statistics for Engineers and Scientists (8th Edition), Prentice Hall, 2006;

[55] – Kaltenbach, H., A Concise Guide To Statistics (1st Edition), Springer, 2011;

[56] – Charkrabti, S., Cox, E., Frank, E., Guting, RH., Han, J., Jiang, X., Kamber, M., Lightstone, SS., Nadeau, TP., Neapolitan, RE., Pyle, D., Refaat, M., Schneider, M., Teorey, TJ., Witten, IH., Data Mining: Know It All, Morgan Kaufmann, 2008;

[57] – Steven, J., Applied multivariate statistics for the social sciences (3rd edition), Lawrence Erlbaum, New Jersey, 1996;

[58] – Vickers, AJ., Parametric versus non – parametric statistics in the analysis of randomized trials with non – normally distributed data, *BMC Medical Research Methodology*, 5:35, 2005;

[59] – Cohen, J., Statistical power analysis for the behavioral sciences (2nd edition), New Yorn, Erlbaum, 1988;

[60] – Hauke, J., Kossowski, T., Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data, *Quaestiones Geographicae*, 30(2): 87-93, 2011;

[61] – Frimodt – Møller, M., Nielsen, AH., Kamper, A., Strandgaard, S., Reproducibility of pulse – wave analysis and pulse – wave velocity determination in chronic kidney disease, *Nephrol Dial Transplant*, 23: 594-600, 2008;

[62] – Horváth, IG., Németh, A., Lenkey, Z., Alessandri, N., Tufano, F., Kis, P., Gaszner, B., Cziráki, A., Invasive validation of a new oscillometric device (Arteriograph) for measuring augmentation index, central blood pressure aortic pulse wave velocity, *J Hypertens*, 28(10):2068-2075, 2010;

[63] –Gorunescu, F., Data Mining: Concepts, Models and Techniques, Springer, 2011;

[64] – Han, J., Kamber, M., Data Mining: Concepts and Techniques (2nd edition), Morgan Kaufmann, Elsevier, 2006;

[65] – Hand, DJ., Statistics and data mining: intersecting disciplines, *ACM SIGKDD Explorations Newsletter*, 1(1):16-19;

[66] – Yoo, I., Alafaireet, P., Marinov, M., Pena-Hernandez, K., Gopidi, R., Chang, JF., Hua, L., Data mining in healthcare and biomedicine: a survey of the literature, *J Med Syst.*, 36(4):2431-2448, 2012;

[67] – Kotsiantis, SB., Supervised Machine Learning: A Review of Classification Techniques, *Informatica*, 31:249-268, 2007;

[68] - Haykin, S., Neural Networks: A Comprehensive Foundation (2nd Edition), Prentice Hall, 1998;

[69] – McCulloch, W., Pitts, W., A logical calculus of the ideas immanent in nervous activity, *Bulletin of Mathematical Biology,* 5(4):115 – 133, 1943;

[70] – Hornik, K., Stinchcombe, M., White, H., Multilayer Feedforward Networks are Universal Approximators, *Neural Networks,* 2:359-366, 1989;

[71] – Sifaoui, A., Abdelkrim, A., Benrejeb, M., On the Use of Neural Network as a Universal Approximator, *IJ – STA,* 2(1): 386 – 399, 2008;

[72] – Breiman, L., Random Forests, *Machine Learning,* 45:5-32, 2001;

[73] – Quinlan, JR., C4.5: Programs for Machine Learning, Morgan Kaufmann Publishers, 1993;

[74] - Cohen, W., Fast effective rule induction, *Machine Learning: Proceedings of the Twelfth International Conference,* Lake Tahoe, California, Morgan Kaufmann, 1995;

[75] – Furnkranz, J., Widmer, G., Incremental reduced error pruning, *In Proceedings of the Eleventh International Conference on Machine Learning,* 70-77, Morgan Kaufmann, 1994;

[76] – Legány, C., Juhász, S., Babos, A., Cluster validity measurement techniques, *Aiked'06 Proceedings of the 5th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases,* 388 – 393, 2006;

[77] – Li, Bing Nan, *et al.,* On an automatic delineator for arterial blood pressure waveforms, *Biomedical Signal Processing and Control,* 5:76-81, 2010;

[78] – Blood Pressure Chart, accessed 28 May, 2012, http://www.disabled-world.com/artman/publish/bloodpressurechart.shtml;

[79] – Van Bortel, LM., Duprez, D., Starmans-Kool, MJ., Safar, ME., Giannattasio, C., Cockcroft, J., Kaiser, DR., Thuillez, C., Clinical applications of arterial stiffness, Task Force III: recommendations for user procedures, *Am J Hypertens,* 15(5):445-452, 2002;

[80] – Luo, X., Yang, Y., Cao., T., Li, Z., Differences in left and right carotid intima – media thickness and the associated risk factors, *Clin. Radiol,* 66(5):393-398, 2011;

[81] – Ferranti, S., Rifai, N., C-reactive protein and cardiovascular disease: a review of risk prediction and interventions, *Clinica Chimica Acta,* (317(1-2):1-15, 2002;

[82] – Scott, J., Pathophysiology and biochemistry of cardiovascular disease, *Current Opinion in Genetics & Development,* 14(3): 271-279, 2004;