



HAL
open science

Dynamical analysis of mutivariate time series for the early detection of syncope during Head-Up tilt test

Nadine Khodor

► **To cite this version:**

Nadine Khodor. Dynamical analysis of mutivariate time series for the early detection of syncope during Head-Up tilt test. Signal and Image processing. Université de Rennes; Université Libanaise. Faculté des Sciences (Beyrouth, Liban), 2014. English. NNT : 2014REN1S123 . tel-01142153

HAL Id: tel-01142153

<https://theses.hal.science/tel-01142153>

Submitted on 14 Apr 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Européenne de Bretagne

En Cotutelle Internationale avec
L'Université Libanaise, Liban

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

Mention : traitement de Signal et Télécommunications

Ecole doctorale Matisse

présentée par

Nadine Khodor

Préparée à l'unité de recherche INSERM, U1099
Laboratoire de Traitement du Signal et de l'Image
UFR ISTIC : Informatique et Electronique

**Dynamical analysis
of multivariate time
series for the early
detection of syncope
during Head-Up tilt
test**

**Thèse soutenue à Beyrouth-Liban
le 22 décembre 2014**

devant le jury composé de :

Catherine MARQUE

PU, Université de Technologie de Compiègne / *Rapporteur*

Mohamad DIAB

Professeur Associé, Université Rafic Hariri / *Rapporteur*

Jacques DUCHENE

PU, Université de Technologie de Troyes / *Examineur*

Ziad ABU-FARAJ

Professeur Associé, Université AUST / *Examineur*

Guy CARRAULT

PU, Université de Rennes 1 / *Directeur de thèse*

Mohamad KHALIL

PU, Université Libanaise / *Directeur de thèse*

Alfredo I HERNÁNDEZ

Directeur de recherche INSERM, HDR / *Co-directeur de thèse*

Hassan AMOUD

EC, Université Libanaise / *Co-directeur de thèse*

Acknowledgements

It would not have been possible to write this doctoral thesis without the help and support of the kind people around me, to only some of whom it is possible to give particular mention here.

I would like to acknowledge AZM and SAADE association, Campus France and the French Lebanese Research Program CEDRE that provided the necessary financial support for this research

I would like to express my deepest sense of gratitude to my supervisors in France: *Guy CARRAULT*, I thank him very much for his perseverance, his infinite patience in all the time of research especially during the writing and the correction of this manuscript. I also thank him for his words of support at critical moments and *Alfredo Hernandez*, for the continuous support of my PhD study and research, for his patience, motivation, enthusiasm, and immense knowledge.

My sincere thanks also go to my supervisors in Lebanon: *Mohamad KHALIL*, his guidance helped me all the time not only in my PhD research but also during my academic studies. *Hassan AMOUD*, the first one who has encouraged me to pursue a doctoral degree, I thank him also to offer his continuous advice and encouragement throughout the course of this thesis.

I thank Dr. Francois CARRE, along with his team members, especially Nathalie VILLE and David MATELOT for the fruitful collaborative efforts that they maintain with my research. Of course, I cannot forget to thanks Philippe L'HOSTIS and Biotrial with whom I have worked during my first years of PhD.

In my daily work, I have been blessed with a friendly and cheerful group of fellow students and members of the following staff , *Azm Center, Campus numérique Francophone de Tripoli* and *Laboratoire Traitement de Signal et de l'Image (LTSI)*. Thanks to all these people. Special thanks to *Lotfi SENHADJI* for welcoming me in his research laboratory (*LTSI*), *Patricia, Muriel, Sozic, Jana EL HAJJ, Rayan NASSER and Youmna SAYOUR* for their logistic support, organizational talents and their kind spirits.

I am thankful to all my friends for their continued help and support: the new ones, for life experiences that we have shared in France and elders, for staying present despite the distance.

Finally, I would like to thank my beloved parents *Maher* and *Fatmeh*, my brothers *Hadi* and *Haroun* and my sisters *Malakeh* and *Rozine* that have given me their unequivocal support throughout, as always, for which my mere expression of thanks likewise does not suffice. I take this opportunity to express the profound gratitude to all my family members especially my grandmother *Malakeh*, my aunt *Sabah* and my uncle *Abdel-Karim* and their children. My last thanks go to the spirit of my dear uncle *Abdelmenhem*.

Résumé en français

La syncope est une perte brusque et transitoire de conscience associée à la perte de tonus postural et suivie d'une récupération spontanée et rapide. Elle est généralement causée par un flux sanguin temporaire insuffisant vers le cerveau. Environ 40% des adultes ont développé une syncope au moins une fois dans leur vie. Seules les statistiques des hôpitaux américains sont disponibles, et ils estiment que 3-5% des visites aux services d'urgence et de 1-6% des hospitalisations sont inscrits en raison de ce trouble. Ces chiffres sont probablement la même en France, mais il n'y a pas de données précises disponibles jusqu'à la date de publication de ce travail.

Bien que la syncope ne soit pas généralement mortelle, elle présente un impact économique sur le système de soins, en plus de son effet sur la vie personnelle, sociale et professionnelle de personnes en souffrant. (St-Jean et al., 2008) ont rapporté que 44% des patients qui ont développé des syncopes récurrentes avaient subi des blessures au cours de la syncope y compris des meurtrissures (32%), des lacérations (9%), des fractures (8%) et de graves blessures internes (10%).

Lorsque la syncope est répétée, le patient peut être invité à réaliser un test d'inclinaison. Ce test, appelé head-up tilt test (HUTT) est une procédure médicale bien identifiée et permet de reproduire le mécanisme de la syncope. Cependant, son problème principal réside dans sa durée, qui ne convient pas aux patients. En outre, le personnel médical est monopolisé pendant près d'une heure, si aucun des symptômes apparaissent. Il importe donc de réduire la durée de ce test et c'est l'objet de la présente étude qui tente de prédire son résultat et d'éviter aux patients de développer une syncope pendant l'essai.

Pendant un HUTT, de nombreuses données sont acquises comme l'électrocardiogramme (ECG) et la pression artérielle. Plusieurs études ont montré leur intérêt pour prédire le résultat de l'HUTT. Cependant, peu d'études ont évalué les méthodes d'analyse non-linéaire de séries temporelles extraites de ces signaux et ont pris en compte leurs dynamiques temporelles. C'est l'objet de ce mémoire de thèse.

Dans un premier temps, l'analyse de la variabilité cardiaque –qui fournit des informations sur la balance sympathique et parasympathique dans le contrôle du rythme cardiaque-, a été étudiée. Les paramètres linéaires classiques et non-linéaires ont été utilisées et comparées au baroréflexe. L'expérimentation, réalisée sur 66 sujets parmi lesquels 35 ont développé une syncope (sujets positifs) et 31 qui n'ont pas développé une

syncope (sujets négatifs) pendant le tilt test, montre une amélioration de la compréhension physiologique du malaise. En utilisant une méthode de classification à noyau et les paramètres non linéaires estimés dès les 15 premières minutes de la position inclinée, une sensibilité de 81% entre les deux groupes et une spécificité de 88% ont pu être observées.

Ensuite, les interactions dynamiques entre des signaux cardiaques (ex. ECG et signal de pression) -qui peuvent jouer un rôle important pour améliorer la compréhension de la régulation cardiovasculaire-, ont été analysées. Deux approches différentes pour caractériser la relation temps-fréquence entre les signaux ont été exploitées. La première approche est basée sur la distribution Pseudo-Wigner-Ville lissée et la seconde utilise un coefficient de corrélation linéaire local dans le domaine temps-fréquence. De ces plans temps-fréquence, deux nouveaux indices, capables de quantifier la relation entre les séries temporelles, ont été estimés. Le premier calcule la densité de la cohérence entre deux signaux dans différentes bandes de fréquence et le deuxième calcule le produit scalaire entre les spectres de temps-fréquence. L'intérêt de prendre en compte ces interactions est souligné et montre, sur la même base de données, une amélioration des performances avec une sensibilité de 85% et une spécificité de 98%.

Bien qu'intéressant, ces résultats nécessitent de débiter le test d'inclinaison pendant 15 min, ce qui reste un test clinique désagréable pour le patient. Ce travail propose donc également une méthode d'analyse de la dynamique de la variabilité de la fréquence cardiaque dans l'espace de phase lors de la phase de repos du test et une modélisation par chaîne de Markov cachées. L'originalité de notre approche est de décrire le plan de phase par un ensemble de paramètres décrivant la densité de passage dans une zone du plan ou bien en exploitant des indices déduits de l'analyse par quantification de la récurrence. Les résultats enregistrés présentent une sensibilité de 95% et une spécificité d'environ 50%.

Cependant, il était aussi important de juger les performances dans chaque fenêtre d'analyse de cette étude en exploitant l'ensemble des modalités enregistrées (signaux ECG et de pression). Une règle de fusion optimale est alors proposée afin de combiner les données provenant de différentes séries temporelles. Cette combinaison a conduit à une amélioration notable des performances dans les différentes fenêtres temporelles. Les performances atteignent 100%, 92,8% de sensibilité et 78%, 100% de spécificité respectivement durant les 5 minutes de repos et les 15 premières minutes de la position inclinée. Ces résultats sont, à notre connaissance, supérieurs aux performances reportées dans la littérature lors de la phase de repos ou du début du test d'inclinaison et cette stratégie d'analyse devient de fait la base du test clinique de détection précoce d'une syncope lors d'un HUTT.

Contents

Acknowledgements.....	<i>ii</i>
Résumé en Français.....	<i>iv</i>
Contents.....	<i>vi</i>
Glossary.....	1
General introduction	3
1 Cardiac electrophysiology and regulation of the cardiovascular system by the autonomic nervous system.....	6
1.1 The heart.....	6
1.1.1 General overview	6
1.1.2 The action potential.....	7
1.1.3 The cardiac conduction system	9
1.1.4 The electrocardiogram (ECG).....	10
1.2 Blood vessels	12
1.2.1 Characteristic of blood vessels	12
1.2.2 Blood pressure.....	12
1.3 Neural control of the cardiovascular system.....	13
1.3.1 The nervous system	13
1.3.2 The autonomic nervous system	14
1.3.2.1 Effect of ANS on the cardiovascular system.....	15
1.3.2.2 Effect on the heart and blood vessels	16
1.3.2.3 Effect on blood pressure.....	17
1.4 Conclusion.....	18
2 Neurally-mediated Syncope and evaluation by Head-Up Tilt Test	19
2.1 Pathophysiology	19
2.2 Epidemiology.....	20

2.3 Classification.....	20
2.4 Reflex syncope (neutrally-mediated syncope)	22
2.5 Head-Up tilt test	24
2.5.1 Introduction.....	24
2.5.2 Methodology	24
2.5.3 Complication and contra-indications	26
2.6 Bibliographic study of syncope prediction.....	26
2.7 Conclusion.....	28
References.....	29
3 Description of the data mining framework applied to early predict syncope	33
3.1 Data selection.....	34
3.1.1 Head-Up Tilt Test protocol.....	34
3.1.2 Subjects	35
3.2 Data processing.....	36
3.2.1 Preprocessing signal	36
3.2.2 QRS complex detection.....	36
3.2.3 Selection and averaging beats	37
3.2.4 ECG features extraction.....	38
3.2.5 Blood pressure features extraction.....	40
3.3 Data mining tools.....	41
3.3.1 Feature selection.....	42
3.3.1.1 The relief method.....	42
3.3.1.2 The sequential forward selection (SFS).....	43
3.3.1.3 'Probe' feature algorithm	43
3.3.2 Classification	45
3.3.2.1 K-nearest neighbor (KNN).....	45
3.3.2.2 Support vector machine (SVM)	46
3.4 Conclusion.....	48
References.....	49

4	Kernel based support vector machine for the early detection of syncope during head-up til test.....	51
4.1	Methods	52
4.1.1	Data processing	52
4.1.1.1	Preprocessing.....	52
4.1.1.2	Feature extraction	54
4.1.2	Data mining	57
4.1.2.1	Statistical analysis.....	57
4.1.2.2	Kernel support vector machine	57
4.1.2.3	Performance evaluation	58
4.2	Results.....	58
4.2.1	Intra-group comparison.....	58
4.2.2	Inter-group comparison.....	59
4.2.3	Classification.....	61
4.3	Discussion	64
4.4	Conclusion.....	67
	Appendix A: Detrended Fluctuation Analysis and Sample Entropy	68
	References.....	69
5	Early syncope detection during head-up tilt test by analyzing interactions between cardio-vascular signals	74
5.1	General overview	75
5.2	Data acquisition and signal processing.....	76
5.3	Time-Frequency analysis method	79
5.3.1	Time-Frequency coherence using SPWD	79
5.3.2	Feature extraction.....	79
5.3.2.1	Distribution of time-frequency coherence over a threshold.....	80
5.3.2.2	The scalar product of the TF spectra	80
5.3.2.3	Secondary indexes	80

5.4 Results.....	80
5.4.1 Frequency band selection and optimization of SPWD parameters.....	81
5.4.2 Classification performances.....	82
5.4.3 Comparison with local linear correlation coefficient.....	83
5.5 Discussion.....	84
5.6 Conclusion.....	86
Appendix B: Data analysis with partial least square regression	88
References.....	90
6 Evaluation of syncope detection coupling a novel phase space analysis algorithm and Hidden Semi-Markov models.....	93
6.1 Analysis of the dynamics of time series	94
6.2 Reconstructed phase space	96
6.2.1 Definition	96
6.2.2 Integration of the time series in RPS.....	97
6.2.2.1 Estimation of the time delay	97
6.2.2.2 choosing the embedding dimension	98
6.2.3 The phase space density	99
6.2.4 Recurrence quantification analysis	100
6.2.4.1 Recurrence plot.....	100
6.2.4.2 Quantitative analysis of the RP	101
6.2.5 Results.....	103
6.2.5.1 RPS's parameters adjustment	103
6.2.5.2 Classification performance	104
6.2.5.3 Performance using PLS	106
6.2.5.4 Performance of other time series in different time-windows....	107
6.2.5.5 Optimal fusion decision rule	109
6.3 Hidden Semi-Markov models	110

6.3.1 Parameter estimation of the HSMM	111
6.3.2 Syncope prediction process	113
6.3.2.1 Feature extraction and preparation.....	113
6.3.2.2 Training and testing the two HSSM models.....	114
6.3.3 Evaluation of the algorithm.....	114
6.4 Conclusion	115
Appendix C: Markov and Hidden Semi-Markov models	117
References.....	121
Conclusion	125
List of publications	128
List of figures.....	129
List of Tables.....	132

Glossary

Abbreviation	Definition of the parameters used in this manuscript
mean_RR	Average of heart rate
std_RR	Standard deviation of RR-interval time series. An estimate of overall HRV
rmssd	Standard deviation of the first derivative of RR-interval time series
pRR50	The percentage of pairs of adjacent normal RR-intervals differing by more than 50 ms
BRS	Non-invasive baroreflex sensitivity computed as the slope between RR-interval and sBP variability
Up _{BRS}	Up sequence computed by the sequence technique
Down _{BRS}	Down sequence computed by the sequence technique
LF	Lower frequency power: A marker of sympathetic activity
HF	High frequency power: A marker of parasympathetic activity
P _{tot}	Total power spectrum
LFnu	Normalized low frequency power
HFnu	Normalized high frequency power
LF/HF	The ratio LF/HF
α_1	Short term correlation properties: The short-range correlation of the detrended RR Intervals
α_2	Long term correlation properties: The long-range correlation of the detrended RR Intervals
SampEn	Sample Entropy: the irregularity in data
$\gamma(t,f)$	Time frequency coherence computed by SPWD. An estimation of the local coupling between two signals in Time Frequency domain

γ_i	Distribution of time-frequency coherence over a threshold in band i
S_i	The scalar product of the TF spectra in band i . An estimation of the similarity between time series in Time Frequency domain
$R^2(t,f)$	Local linear correlation coefficient: represents the local coupling between two signals in Time Frequency domain
R_i	Local linear correlation coefficient in band i
$R_e R_n$	Recurrence Rate: the percentage of recurrence points in a Recurrence plot (RP)
DET	Determinism: the percentage of recurrence points which form diagonal lines in RP
LAM	Laminarity: the percentage of recurrence points which form vertical lines in RP
$\langle L \rangle$	The average length of the diagonal lines
TT	Trapping Time: The average length of the vertical lines
L_{max}	The length of the longest diagonal line
V_{max}	The length of the longest vertical line
$ENTR$	The shanon entropy of the probability distribution of the diagonal line length
$RPDE$	Recurrence period density entropy

General Introduction

Syncope is an abrupt and transient loss of consciousness associated with loss of postural tone and followed by rapid spontaneous recovery. It is usually caused by temporary insufficient blood flow to the brain. Around 40% of adult population have experienced syncopal episode at least once in their life. Only American hospital statistics are available, and they estimate that 3-5% of emergency department visits and 1-6% of hospital admissions are enrolled due to this disorder. These numbers are probably the same in France but there is no precise data available until the date of publishing this work.

Although syncope is not generally life threatening, nevertheless it has economic relevance on the healthcare system, in addition to its considerable impact on the personal, social and professional lives of afflicted person. (St-Jean et al., 2008)¹ have reported that 44% of patients with recurrent syncopes had sustained injuries during the syncopal attacks including bruising (32%), lacerations (9%), fractures (8%) and serious internal injuries (10%). In addition recurrent episodes can be physiologically devastating and can be a premonitory sign of cardiac arrest especially in patients with organic heart disease. All these difficulties lead to decrease the quality of life of patients.

When the syncope is repeated, the patient can be asked to realize Head-Up Tilt Test (HUTT). This test is a well-known medical procedure used to diagnose syncope. It aims to recreate the conditions in which the patient experiences the symptoms of syncope. The test begin with a supine period (about 12-15 min) where the patient must lie on the examination table in a horizontal position; in this period the blood pressure and heart rate are controlled to ensure their stability. Following the stabilization phase and under the action of an electric motor, the table is tilted at an angle between 60° and 80° for duration up to 45 min (It takes around 10 sec to pass from 0° to 80°). Thus the abrupt transition from lying to standing position may trigger syncope, in this case the test is considered as positive and the table is returned to its horizontal position, otherwise the test is considered as negative.

By reproducing the syncope's mechanism, this test can provide information that helps to adapt the treatment of patients. However, the main problem of this test is in its long duration, which is unsuitable for patients especially who have physical weakness

¹ St-Jean, K., Kus, T., Dupuis, G., Lévesque, K., Thibault, B., Guerra, P.G., Nadeau, R., D'Antono, B., 2008. Quality of life in patients with recurrent vasovagal or unexplained syncope: Influence of sex, syncope type and illness representations. *Appl. Res. Qual. Life* 3, 235–249. doi:10.1007/s11482-009-9058-x

and cannot withstand to the end of the test. Additionally, medical staff is monopolized for almost an hour if no symptoms appear. Therefore, minimizing this duration helps to optimize the time of clinical practitioners, provide more comfortable conditions for patients and will reduce costs. For this purpose, this study was realized; it attempts to early predict the outcome of HUTT and to avoid patients to experience syncope during the test. Within this framework, this study aims also to help the doctors to obtain a deduction of the results of the examination process before its ending as: the presence or the absence of symptoms related to the development of syncope.

During HUTT, numerous measurements are acquired such as electrocardiogram (ECG) and blood pressure signals, which could be used to elaborate predictive criterion for the final result. Several studies have been realized to exploit the evolution of these signals during the test allowing to search methods for early predict the results. However, few studies have evaluated the non-linear analysis methods of time series extracted from these signals and their inter-relationship and took in consideration their temporal dynamics.

The objective of the work realized in this manuscript is to propose and evaluate signal processing methods that exploit the dynamics time series extracted from cardiovascular signals, especially the ECG and pressure signals, to develop new tools for the early prediction of syncope occurred during HUTT and to better understand the mechanisms behind this fainting. The methodological aspects investigated in this study are decomposed in six chapters.

The first chapter constitutes a general overview of the physiological basis of the cardiac activity and its relation with autonomic nervous system to better understand the origin of the syncope. After this brief review, the acquisition of the electrical activity of the heart is explained as well as the origin and the indications of elementary waves constituting the ECG signal.

Chapter 2 focuses on the positioning of the problem and on the definition of the syncope, which represents the main subject of this study. First, the pathophysiology, the epidemiology and the classification of syncope are detailed. Then, the standard procedure, HUTT, used to reproduce this clinical event in patients with suspected syncope is described. To end with a bibliographical review on the different methods already proposed and applied to predict syncope. This part highlights the difficulty of this task and we will see that the non linear characterization and the dynamic properties of prominence features extracted from cardiovascular signals, and which can provide

pertinent information, are not totally explored, which led us to formalize and propose methods to investigate this unexplored part.

The third chapter presents the general data-mining framework implemented in our study and explains in details the constitution of different steps of this framework; it describes the applied HUTT protocol and the database used to evaluate our algorithms. After that, a segmentation algorithm based on wavelet decomposition is reported, as well as the different feature selection algorithms and classification methods proposed to export our results. Indeed, several feature selection and classification tools are applied to compare between them in solving our problem.

Chapter 4 takes us to the core of the considered problem, and studied the heart rate variability (HRV) which provides information about the sympathetic and parasympathetic balance that control cardiac rhythm. Conventional linear and non-linear parameters were extracted from HRV as well as the baroreflex syncope in different time windows. The capability of these parameters to differentiate between negative and positive test groups were evaluated using the data mining tools described in the previous chapter.

Then the dynamic interactions between cardiac signals (ECG and blood pressure), which can play an important role in improving the understanding of the cardiovascular regulation, were analyzed in chapter 5. Two different approaches to characterize the time- frequency (TF) relation between signals have been exploited. The first approach is based on the smoothed pseudo-wigner ville distribution and the second one, uses a local linear correlation coefficient in the time frequency domain. In this TF plan, two new indexes able to quantify the relationship between the time series were estimated. The first computes the density of the coherence between two signals in different frequency bands and the second calculates the dot product between the TF spectra.

Although interesting, these findings need to start tilt test for 15 minutes, which is still in the clinical domain uncomfortable for the patient. This work has therefore proposed a method to analyze the dynamics of the heart rate variability whether by the phase space during the supine period of the test or by the hidden semi-markov model (chapter 6).

Chapter 1

Cardiac electrophysiology and regulation of the cardiovascular system by the autonomic nervous system

The cardiovascular system is composed mainly of the heart and the blood vessels (arteries, veins, blood capillaries). Its basic function is to provide a continuous flow of blood to organs and cellular tissues of the body in order to a) distribute oxygen and nutrients, b) eliminate waste and carbon dioxide generated during their activity and c) transport hormones produced by endocrine glands to the receivers. This chapter aims to address the physiological basis of the cardiac activity and its relation with the autonomic nervous system, to understand the origin of vasovagal syncope. It draws heavily from previous PhDs realized in our laboratory [Dumont, on 2008, Wong on 2004, Hernández, on 2000]. After the description of the heart and its fundamental basis (the action potential, the cardiac conduction system and the Electrocardiogram signal), the main characteristics of the blood vessels are described. Then, the neural control of the cardiovascular system is briefly overviewed.

1.1. The heart

1.1.1. *General overview*

The heart is a muscular organ, which insures the blood circulation in veins and arteries. It is situated within the chest cavity and is slightly moved to the left. The heart of an adult weighs 300 to 350 grams, it pumps an average of 70 times per minute, propelling each day about 7,000 liters of blood in the cardiovascular system.

The heart has four chambers (figure 1.1), two upper atria, the receiving chambers, and two lower ventricles, the discharging chambers. The right atrium receives deoxygenated blood from the body and the left atrium receives oxygenated blood from the lungs. When they contract, the blood is pushed into the ventricles, which pump to

propel the blood to the lungs and the rest of the body. The inter-atrial septum separates the atria and the inter-ventricular septum separates the ventricles. Ventricles are separated from atria by means of atrio-ventricular valves formed by ailerons of connective tissue which provide in the normal state a unidirectional path of the blood: tricuspid valve for the "right heart "; mitral valve for the "left heart ". Then, other valves ensure the communication between the right ventricle and the pulmonary artery (pulmonary valve consists of three sigmoid valves) and between the left ventricle and the aorta (aortic valve). These valves prevent backflow into the ventricles, once the blood is deported to the pulmonary artery and to the aorta.

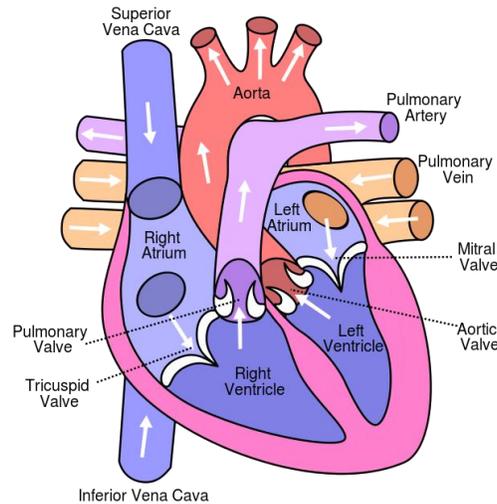


FIG. 1.1 – Anatomical structure of the heart [CC²](#).

The mechanical cardiac activity described above is initiated and controlled by a cardiac electrical activity, which is generated at the cellular level and is known as the “action potential”.

1.1.2. The action potential

Most cardiac cells are electrically excitable. In the cell’s rest period, the inside of the cellular membrane is negatively charged compared to the outside. This difference of potential is the reference situation. It depends on ionic concentrations in the intracellular and extracellular environments; its value is close to -90 mV for ventricular cells.

When an excitable cardiac cell is electrically activated, it generates a reverse potential difference (the inside of the cell rapidly becomes positive compared with the outside). This phenomenon is called cellular "depolarization". The return of the cardiac

² CC: Picture with Creative Common copyright..

cell to its resting state is named "repolarization" (figure 1.2). Potential differences measured between the intracellular and extracellular environments during these two phenomena define the **action potential (AP)**, which is characterized by five different phases (figure 1.3):

- **Phase 0 or fast depolarization:** after an electrical excitation over the cell activation threshold, a fast flow of Na^+ ions goes into the cell and quickly inverts its polarity.
- **Phase 1 or the initial state of repolarization:** it is characterized by a rapid repolarization of short duration, due to the inactivation of Na^+ channels and the outgoing flow of K^+ ions.
- **Phase 2 or plateau stage:** where the rate of repolarization is slowed by the influx of Ca^{2+} ions into the cell. The Ca^{2+} ions enter the cell slower than the Na^+ ions and help prevent the cell from repolarizing too quickly.
- **Phase 3 or the later stage of repolarization:** It is characterized by the closure of specific ionic channels which returns the cell to the reference potential. It is dependent on the increased flow of potassium out of the cell while the potential of the cell tends towards its threshold of rest.
- **Phase 4:** It corresponds to the rest potential, where the cell is more easily excitable.

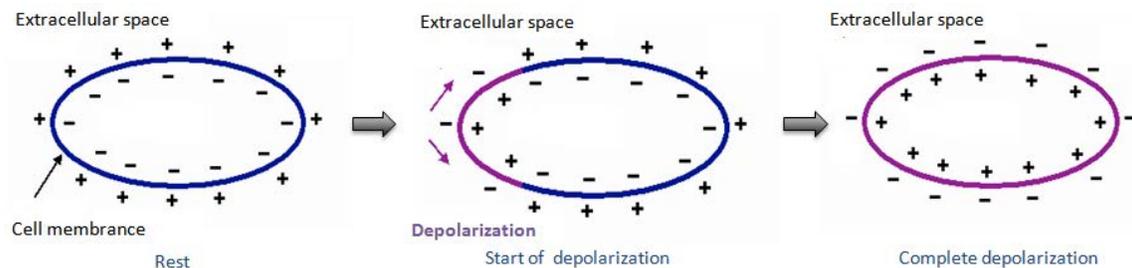


FIG. 1.2 – Depolarization-repolarization cycle

From the beginning of phase 1 until nearly the half of phase 3, each cell is in an absolute refractory period (ARP), during which it is impossible to evoke another action potential, followed by a relative refractory period (RRP), where a stronger-than-usual stimulus is required to generate an AP. These two refractory periods are caused by changes in the state of sodium and potassium channel molecules. When closing after an action potential, sodium channels enter an "inactivated" state, and they cannot be opened regardless of the membrane potential—this gives rise to the absolute refractory period. Even after a sufficient number of sodium channels transitioned back to their resting state, a fraction of potassium channels remains open, making difficult but

possible for depolarization to occur, and thereby giving rise to the relative refractory period.

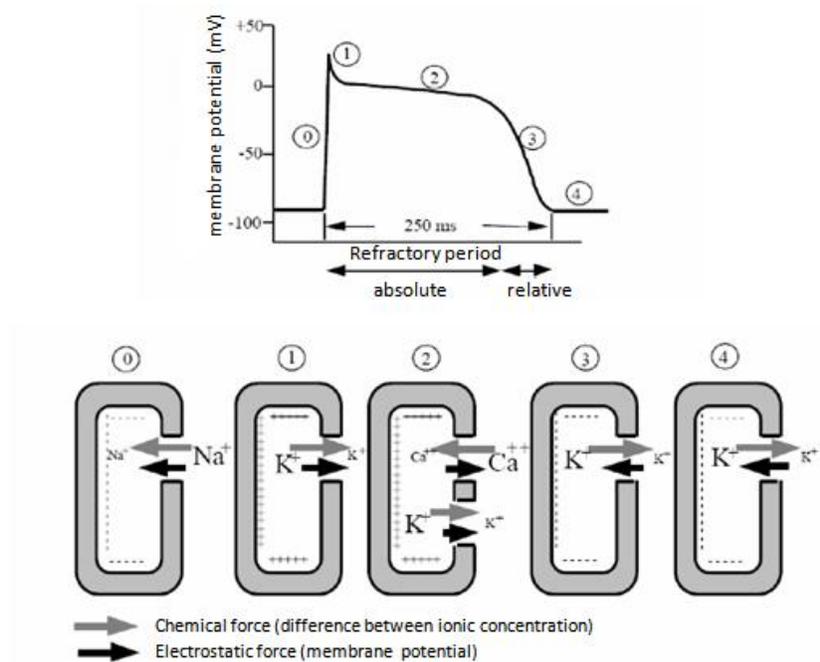


FIG. 1.3 – The action potential and the corresponding ionic exchange.

1.1.3. The cardiac conduction system

Even isolated, the heart continues to work and to contract rhythmically: we say that the heart is endowed with automation. This autonomous functioning is ensured by a specific excitation and conduction system illustrated in figure 1.4 and described below.

Myocardial contraction is caused by the propagation of an electrical pulse along the cardiac muscle fibers induced by the depolarization of muscle cells. Indeed, the heart contains an intrinsic network of conductive cells which produce and propagate electrical pulses, as well as cells that respond to these pulses by a contraction. During a normal cardiac activity, the electrical stimulation of the myocardium arises from the sinoatrial (SA) node, the natural pacemaker of the heart. After passing through the atrium, this electrical stimulation passes through the atrioventricular node (AV node) before joining the ventricles via the nodal distribution network which is the bundle of His and the Purkinje fiber.

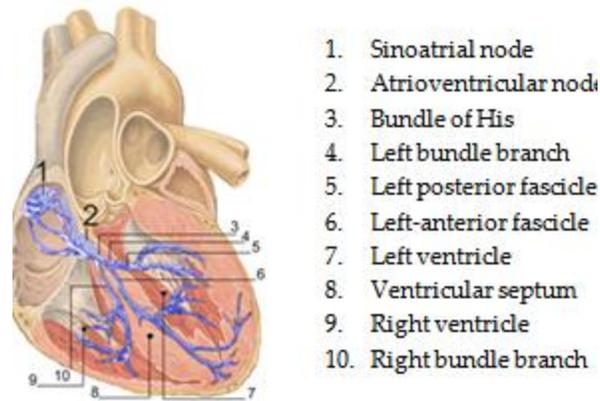


FIG. 1.4 – The cardiac conduction system [CC](#)

1.1.4. *The electrocardiogram (ECG)*

In 1880, E. Marey and A. Waller showed that the electric activity of the heart, discovered some years earlier, could be observed from the skin; and by 1890, Willem Einthoven, realized the first electrocardiographic recording. The current measured by electrodes on the trunk of the patient put in movement a thin silver wire stretched between the poles of a big magnet; these deflections were recorded on photographic paper which unfolded in front of a beam of light. So the electrocardiogram was born at the dawn of the 20th century.

The flow direction and the amplitude of the electric currents generated by the depolarization and repolarization processes of the myocardial cells can all be detected by electrodes placed on the surface of the thorax (Figure 1.5). By examining an ECG, we recognize (figure 1.6):

- **The isoelectric line:** a line of zero potential.
- **The atrial depolarization:** represented on the ECG by the **P** wave. This wave is characterized in the spectral domain by a low-frequency component of low-energy.
- **The atrial repolarization:** represented by the **Ta** wave, it is in opposite direction of the **P** wave. Generally **Ta** wave is not visible in the ECG as it coincides with the **QRS** complex of greater amplitude.
- **The ventricular depolarization:** represented by the deflection of larger amplitude of the ECG, the **QRS** complex. It constitutes by three consecutive waves: **Q**, **R** and **S** waves.
- **The Ventricular repolarization:** reflected by the asymmetric **T** wave with the ascending slope is lower than the downward slope.

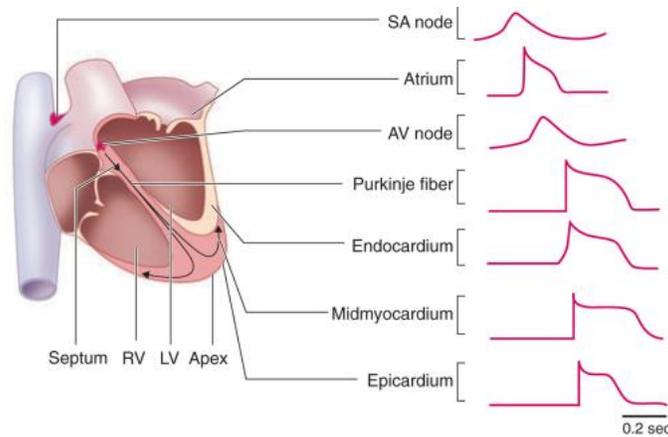


FIG. 1.5 - Propagation of the action potential in the heart (Nerbonne and Kass, 2005) .

In addition to the waveform, a heartbeat is characterized by a number of segments and intervals (Figure 1.6):

- **PR interval:** it is measured from the beginning of the **P** wave to the beginning of the **QRS** complex. This interval represents the depolarization of the atria and the AV node. Its normal duration is between 0.14 and 0.2 s;
- **The PR segment:** it is the time between the end of the **P** wave and the beginning of the **QRS** complex. It represents the time of transmission of the depolarization wave by the AV node;
- **The ST segment:** it is between the end of the **QRS** complex (or point J) and the beginning of the ascending phase of the **T** wave. This segment represents the time during which all the myocardial cells are depolarized (phase 2) so, in the normal case should be isoelectric. Otherwise, the amplitude level and the slope of this segment are indicators of myocardial ischemic state;
- **QT interval:** the time between the beginning of the **QRS** complex and the end of the **T** wave. It is an indication of the length of the ventricular depolarization and repolarization phases. Its duration varies according to the heart rate between 0.3 and 0.38 s.

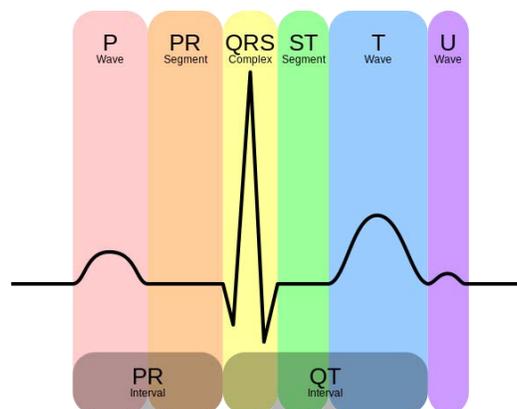


FIG. 1.6 - Waves, intervals and segments of a heart beat in ECG [cc](#).

1.2. Blood vessels

1.2.1. Characteristics of blood vessels

The vascular system includes several types of vessels that together provide the transportation of the blood, from the heart to the organs and conversely:

- **The arteries:** They present a muscular and elastic thick wall. They can branch out into smaller vessels, arterioles.
- **The Veins:** They have thin walls, which cannot resist a strong pressure as the arteries. They can branch into smaller vessels, venules.
- **The Capillaries:** they are tiny, extremely thin-walled vessels that act as a bridge between arteries and veins.

Table 1.1 summarizes the characteristics and the functions of the different types of blood vessels.

TAB 1.1 - Characteristics and functions of blood vessels.

Type	Blood pressure	Direction of blood flow	Function
Arteries	Strong	heart--> body	Carry blood enriched in oxygen and nutrients
Veins	Rather small	body --> heart	Carry blood containing carbon dioxide
Capillaries	Very low	connect arterioles and venules	Exchange area between the blood and organs

1.2.2. Blood pressure

The blood pressure reflects the force exerted by the blood on the walls of arteries. It is usually expressed by two values in most clinical applications:

- The systolic pressure measured during the ventricular systole (contraction) of the heart when blood is ejected with the highest power;
- The diastolic pressure, measured during the ventricular diastole (relaxation).

However, current data acquisition tools allow for the acquisition of a continuous blood pressure signal in a non-invasive manner. The exchanges of the arteries, arterioles and veins contribute significantly to blood pressure changes. In general, blood pressure increases by vasoconstriction and decreases by vasodilatation. This blood pressure is the controlled variable of the circulation. It must be maintained at an optimal value, sufficient to ensure effective organ perfusion. It depends on the cardiac output (Q') (which itself depends on heart rate and cardiac contractility), and the vascular peripheral resistances (R_p) (eq.1.1). These resistances depend mainly on the diameter of arterial vessels:

$$Pa=Q' * R_p \quad (1.1)$$

1.3. Neural control of the cardiovascular system

Cardiac beats and their rhythm are determined by the intrinsic activity of cardiomyocytes (contractile cells of the cardiac muscle tissue) of the sinoatrial node. Indeed, the cardiomyocytes are spontaneously excitable; their rhythmic depolarization and repolarization are independent of the nervous system. However, the heart rate can be quickly modified by the activity of neurons that innervate the heart. Also, the autonomic nervous system (ANS) controls the frequency of discharge of the sinus node, the force of the ventricular muscular fibers and the resistance of vessels against the blood supply of each beat.

1.3.1. The nervous system

The nervous system (NS) includes billions of nerve cells called neurons. These cells communicate with each other through synapses allowing the passage of nerve impulse. The transmission occurs through neurotransmitters (epinephrine known as adrenalin, acetylcholine etc.) having excitatory or inhibitory effects. The NS is divided into two subsystems (figure 1.7):

- **The Central Nervous System (CNS):** composed of the brain and spinal cord. It transmits sensory information from the periphery to the brain, and responsible for the supervision and the coordination of the entire NS.
- **The Peripheral Nervous System (PNS):** formed of glands and nerves, it circulates information between the organs and the CNS, and it carries out the motor commands of the latter.

The PNS system is divided into:

- **The somatic nervous system (SNS):** controls skeletal muscle as well as external sensory organs such as the skin. This system is known by voluntary because the responses can be controlled consciously. However, Reflex reactions of skeletal muscle are an exception. These are involuntary reactions to external stimuli.
- **The autonomic nervous system (ANS):** used to regulate various body functions (cardiovascular, pulmonary...).

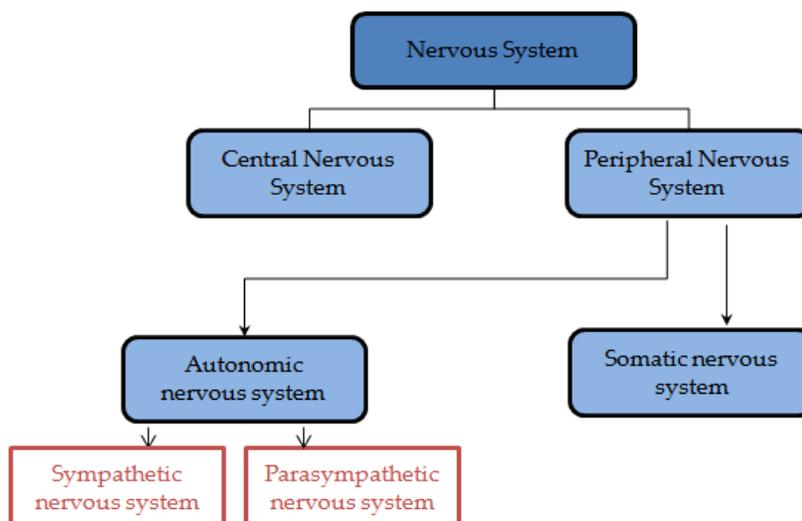


FIG. 1.7 - The nervous system.

1.3.2. *The autonomic nervous system*

The ANS plays a primary role in the maintaining of the homeostasis³ through the regulation of the internal functions of the body. The efferent signals are transmitted to the various organs of the body by two large subdivisions, which are sympathetic and parasympathetic divisions (Figure 1.7). The sympathetic and parasympathetic nerves secrete one of the neurotransmitters: acetylcholine or norepinephrine. Those liberating acetylcholine are called cholinergic fibers and others who release norepinephrine are named adrenergic fibers.

Figure 1.8 represents a general anatomical scheme of both branches of ANS. All the preganglionic neurons are cholinergic. Postganglionic neurons of the parasympathetic nervous system are also cholinergic and most of the sympathetic postganglionic neurons are adrenergic. Before the acetylcholine or norepinephrine provokes their effect on the target organ, these mediators have to settle on specific receivers, which are of two types:

- **Cholinergic receptors (stimulated by the binding of acetylcholine):** they include two types of receptors, muscarinic and nicotinic. The muscarinic receptors are located in all effector cells stimulated by the postganglionic neurons of the parasympathetic nervous system as well as those stimulated by the postganglionic cholinergic neurons of the sympathetic nervous system. Nicotinic receptors are located in the ganglionic synapses of the sympathetic and parasympathetic systems.

³ Homeostasis is a process that maintains the stability of the human body's internal environment in response to changes in external conditions.

- **Adrenergic receptors (stimulated by the binding of norepinephrine):** These include two types of receptors: the alpha (α) and the beta (β) receptors. The organs which react with norepinephrine present one of these receptors or both. In general, the connection of the norepinephrine to α receptor has an excitatory effect, while its binding to β receptors has an inhibitory effect. There are some notable exceptions such as the binding of norepinephrine to β receptors of the heart muscle, which stimulates the activity of the heart.

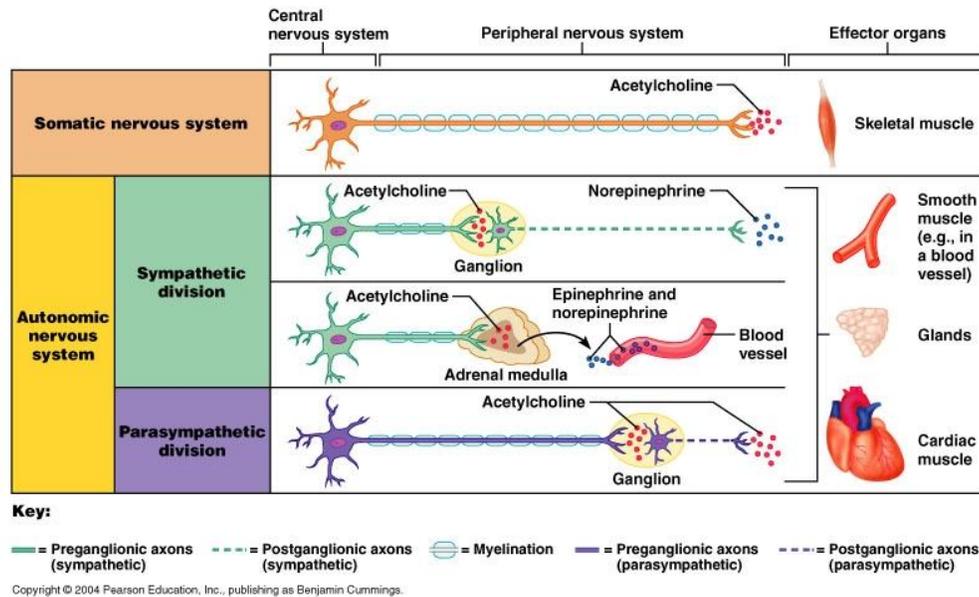


FIG. 1.8 - General anatomical scheme of ANS⁴.

1.3.2.1. Effect of ANS on the cardiovascular system

Table 1.2 shows the effects caused by the stimulation of the parasympathetic or sympathetic nerves on various cardiovisceral functions of the body. From this table, we can see that sympathetic stimulation causes excitatory effects on some organs but inhibitory effects on others. Similarly, parasympathetic stimulation causes excitation of certain organs and inhibiting others. Also, when sympathetic stimulation excites a particular organ, parasympathetic stimulation may inhibit this latter, demonstrating that the two systems have opposing actions to ensure the proper functioning of the body.

In general, the sympathetic nervous system mobilizes the body during stress such as exercise, while the parasympathetic nervous system allows us to relax while our body performs routine tasks and saves its energy.

⁴ <http://www.austincc.edu/rfofi/NursingRvw/PhysText/PNSefferent.html>

TAB 1.2 - Effect of sympathetic and parasympathetic activation

Organ	Sympathetic stimulation	Parasympathetic stimulation
Heart	Increase heart rate	Decrease heart rate
Heart	Increase contractile force	Decrease contractile force
Coronary artery	Constriction (α), dilatation (β_2)	Dilatation
Vessels	Constriction	No effect
Arteriole	Constriction	No effect
Blood pressure	Short-term increasing	Short-term decreasing

1.3.2.2. Effect on the heart and blood vessels

- **Effect of the parasympathetic nervous system**

Its principal neurotransmitter is the acetylcholine, which is basically for muscarinic receptors. The stimulation of cardiac muscarinic receptors (subtype M₂), slows the conduction of nodal tissue, whereas the stimulation of vascular muscarinic receptors (subtype M₃) causes vasodilation. Parasympathetic stimulation decreases the rate of the sinus node and thus depresses the excitability of the atrioventricular junction fibers, slowing the transmission of nerve impulses to the ventricles. Intense vagal stimulation may even completely inhibit the sinus influx or completely block the transmission at the atrioventricular junction.

- **Effect of sympathetic nervous system**

Its principal neurotransmitter is the norepinephrine, its action is reinforced by the adrenaline. These catecholamines may act on adrenergic receptors that are located mainly in the vascular walls. Its stimulation induces vasoconstriction, thus a rise in blood pressure by increasing peripheral resistance. The stimulation of vascular β_2 receptors is accompanied by a vasodilatation effect.

The effects of sympathetic stimulation oppose almost with those of parasympathetic stimulation. It increases the discharge rate of the sinus node and stimulates ubiquitously the conduction of excitability in the heart. In addition, it increases the contractile force of the heart muscle in atrial and ventricular level. Briefly, the sympathetic NS increases at all levels the cardiac activation. Maximal stimulation can almost triple heart rate and double the strength of the heart contraction.

1.3.2.3. Effect on blood pressure

The cardiovascular system controlled by the autonomic nervous system is a closed loop system. In this short-term loop of the regulation of blood pressure, known as the baroreflex arc, the autonomic nervous system is the controller, the arterial baroreceptor is the main sensor and the heart is the main actuator (Figure 1.9). The nervous messages coming from baroreceptors (BR) are transmitted to the autonomic nervous centers from which it backs to the cardiovascular sympathetic and parasympathetic nerves. These act on heart rate, contractility of myocardial fibers and systematic blood flow resistance in the blood vessels. The controller allows the regulation of blood pressure, by controlling cardiac output and peripheral resistance.

At the level of the cardiac tissue, the exerted effect will be essentially a modulation of the heart rate and the contractility. The relation between arterial pressure and heart rate is a sigmoidal curve, centered on the balance blood pressure. Near this point, the system operates with a maximal sensitivity. When we go away from this point, the system becomes less effective and we can, in extreme cases, end up at two flat responses: the plateau of minimal heart rate beyond which any increase in blood pressure will not result additional bradycardia and conversely the plateau of maximal heart rate where any reduction in blood pressure is not accompanied by a more pronounced tachycardia.

At the vascular level, the ANS is responsible for the modulation of the peripheral resistances: vasodilatation in case of high blood pressure and conversely vasoconstriction when it drops in blood pressure. These two effects (vasodilatation, vasoconstriction) are mainly induced by changes in sympathetic tone.

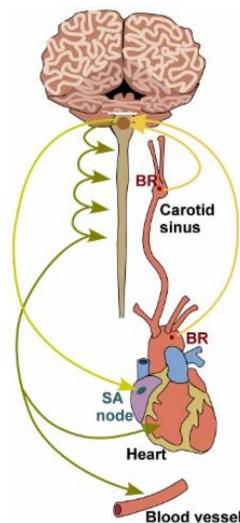


FIG. 1.9- The baroreflex (Mahdi et al., 2013)

1.4. Conclusion

In this chapter, we briefly described the cardiovascular system and its regulation by the ANS. Abnormal behavior of the autonomic nervous system can provoke serious pathologies involving the main functions of the cardiovascular system. ANS function can be indirectly observed by changes in the ECG or blood pressure signals (amplitude, wave duration, interval duration inter-beat directions,...). The following chapter discusses one such pathology, vasovagal syncope, which is the main objective of our study.

References

- Mahdi, A., Sturdy, J., Ottesen, J.T., Olufsen, M.S., 2013. Modeling the Afferent Dynamics of the Baroreflex Control System. *PLoS Comput Biol* 9, e1003384. doi:10.1371/journal.pcbi.1003384
- Nerbonne, J.M., Kass, R.S., 2005. Molecular physiology of cardiac repolarization. *Physiol. Rev.* 85, 1205–1253. doi:10.1152/physrev.00002.2005

Chapter 2

Neurally-mediated Syncope and Evaluation by Head-Up Tilt Test

This chapter presents a synthetic position of the problem and defines vasovagal syncope, which represents the subject of this study. Vasovagal syncope is a symptom, defined as a transient loss of consciousness characterized by rapid onset, short duration, self-limiting and accompanied by a loss of postural tone with a rapid return to normal state of consciousness (Moya et al., 2009). In this chapter, the pathophysiology, the epidemiology and the classification of syncope are discussed. Final sections are dedicated to present the well known medical procedure used to diagnose syncope (Head-up tilt test), and previous works existing in the literature that have focused on predicting syncope, highlighting the complexity of this problem.

2.1. Pathophysiology

The underlying mechanism of syncope is a decrease in the global and transient cerebral blood flow. However, cerebral blood flow depends on the systemic arterial pressure. Any element that decreases whether the cardiac output (*i.e.* venous retention, brady-/tachycardia, vagal reflex, etc.), or the vascular resistance (*i.e.* Thermal stress, antihypertensive drugs, vasodilation, etc.) can decrease the systematic blood pressure, and can provoke syncope. Hypocapnia may also increase the cerebral vascular resistance and lead to reduction in cerebral perfusion.

Generally, loss of consciousness occurs after a stop of 6 to 8 seconds in cerebral blood flow (Kabat H and Anderson JP, 1943) or in a decrease in systolic blood pressure below 60 mmHg (Sheldon and Killam, 1992). A number of compensatory mechanisms are essential to maintain an adequate cerebral blood flow:

1. Cerebral vascular resistance is adapted to the partial pressures of CO₂ and O₂ in the bloodstream.
2. Control of cardiac output and systemic vascular resistance through the baroreflex.
3. Maintaining blood volume and arteriolar vasoconstriction by activating the renin-angiotensin-aldosterone system.

2.2. Epidemiology

Syncope is the cause of almost 3% of all emergency department visits and 1% of hospitalizations (Martin et al., 1984). Although syncope occur at any age, the probability increases with age and in industrialized and post-industrialized countries where the number of people who may develop syncope has strongly grown (Benditt, 2006).

According to some sources (Soteriades et al., 2002), up to 30% of adults are likely to have a syncopal episode at some point in their lives. The age-related variations are still evident and two peaks of incidence are discernible, one in 15-19 years old and one in the 60-70 years old. The most common age at which patients have a first syncope of vasovagal origin is the 13 years old (Sheldon et al., 2006), but syncope increases with advancing age (Savage et al. 1985) (Soteriades et al., 2002) and is generally associated with secondary morbidity to accidental falls (Kapoor, 1987). On the other hand, some studies suggest that women are more likely to develop a syncopal episode (Parry and Tan, 2010), while other studies such as Framingham study, suggested a more balanced impact on sex (Soteriades et al. 2002).

2.3. Classification

Syncope must be differentiated from other « non-syncopal » states associated with real or transient loss of consciousness according to the schema shown in (Figure 2.1). As defined before, blood pressure is determined by the cardiac output and total peripheral vascular resistance. A fall in either one of them can cause syncope, but a combination of both mechanisms is often present, although their relative contributions vary considerably. Figure 2.2 shows a pathophysiological classification based on the main causes of syncope (Moya et al., 2009), with low blood pressure / global cerebral hypoperfusion in the center, concomitant with low or inadequate peripheral resistance and low cardiac output. An inadequate or low peripheral resistance may be caused by:

- Improper reflex activity causing vasodilation and bradycardia manifested by reflex syncope of type vasodepressive, cardio-inhibitory or mixed (shown in red part of Figure 2.2).

- Functional and structural deficiencies of the autonomic nervous system induced by a drug or primary or secondary dysautonomia.

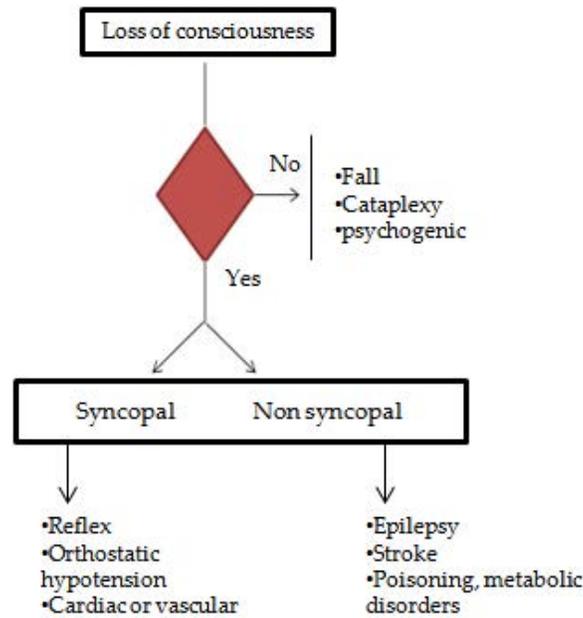


FIG. 2.1 - Classification of brief loss of consciousness.

Factors leading to a transient reduction in cardiac flow are:

- A reflex causing bradycardia known as cardio-inhibitory reflex syncope.
- Cardiovascular problems such as arrhythmias or structural disease including pulmonary embolism, and hypertension (the blue part of figure 2.2).
- Inadequate venous return, due to volume depletion or venous pooling presented in the yellow part of figure 2.2 such as syncope secondary to orthostatic hypotension (OH orthostatic hypotension).

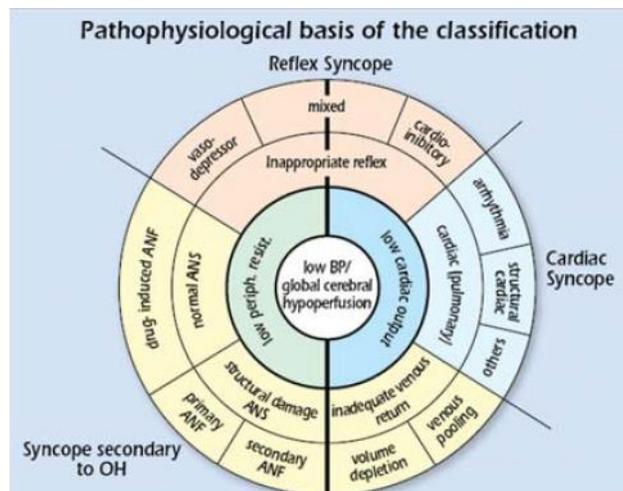


FIG. 2.2 - Pathophysiological basis classification of syncope (Moya et al., 2009)

The prevalence of syncope's causes differs from a study to another because it depends on several factors (age, geographical factors, clinical conditions in which the patient is evaluated, etc.) but some general observations are however possible:

- The Reflex syncope is the most common cause of syncope in any context.
- The Syncope related to cardiovascular disease is the second most common cause.
- In Patients < 40 years old, orthostatic hypotension is a rare cause of syncope; but it is common in elderly patients.

In this manuscript, we will focus on the most common type of syncope: Reflex syncope.

2.4. Reflex Syncope (neurally-mediated syncope)

Neurally-mediated syncope is a synonym for reflex syncope which emphasizes the role of the nervous system in syncope. The term neurally-mediated syncope, is widely used in the version of the recommendation European reference, has been translated automatically by reflex syncope. Reflex syncope refers to a heterogeneous group of conditions in which cardiovascular reflexes normally useful in controlling the circulation become intermittently inadequate in response to a trigger, which causes vasodilation and / or a drop in arterial pressure and global cerebral perfusion (J Gert van Dijk, 2008).

Reflex syncope may also be classified according to its trigger. Note that this is a simplification because different mechanisms may be present in the context of a particular situation (urination, defecation):

- **Vasovagal syncope (VVS):** It is mediated by an emotion or a postural change. It is usually preceded by prodromal symptoms of autonomic activation (sweating, pallor, nausea). In VVS, following a trigger stimulus, i.e. moving to standing position or prolonged standing position, an irregularity of the posture change induces an abnormal decrease in venous return to the heart. The sympathetic activation of the heart (which causes an increase in the heart rate and cardiac contractility) in this context, drops this discharge: the heart is a "vacuum pump". These strong "empty" contractions can trigger the Bezold-Jarisch reflex (Mosqueda-Garcia et al., 2000) which involves a massive parasympathetic activation that decreases blood pressure and heart rate. This mismatch would be responsible for cerebral ischemia and syncope (Benditt et al., 2004), (Bloomfield et al., 1999).
- **Situational syncope:** It occurs in a particular context (urination, defecation, coughing, swallowing) (Kapoor et al, 1985.) (Kapoor et al., 1986). To cite some examples, the micturition syncope occurs during or immediately after urination.

They are secondary to over-stimulation of baroreceptors that are in the bladder. Among the predisposing factors include alcohol, fasting, fatigue or an infectious condition.

- **Carotid sinus syndrome:** In its rare spontaneous form, it is triggered by mechanical manipulation of the carotid sinus. In the most common form no mechanical trigger is found and it is diagnosed by carotid sinus massage (Tea et al., 1996).
- **Atypical form:** This term is used to describe situations in which the reflex syncope occurs with uncertain or even apparently absent triggers. The diagnosis is then based on the exclusion of other causes of syncope (absence of structural heart disease) and on the reproduction of similar symptoms with orthostatic tilt test.

Reflex syncope can be also classified according to the most involved efferent pathway (sympathetic or parasympathetic), (figure 2.3):

- **Vasodepressor type:** also called VASIS III, according to the international study of vasovagal syncope (VASIS: Vasovagal syncope international study) (Brignole et al., 2000). During the tilt test, when syncope occurs, heart rate (HR) suddenly drops over 10% and a hypotension appears (Moya et al., 2009).
- **Cardio-inhibitory type:** or VASIS II, when bradycardia predominates and precedes a drop in blood pressure.
- **Mixed type:** or VASIS I, when the two mechanisms are simultaneously present.

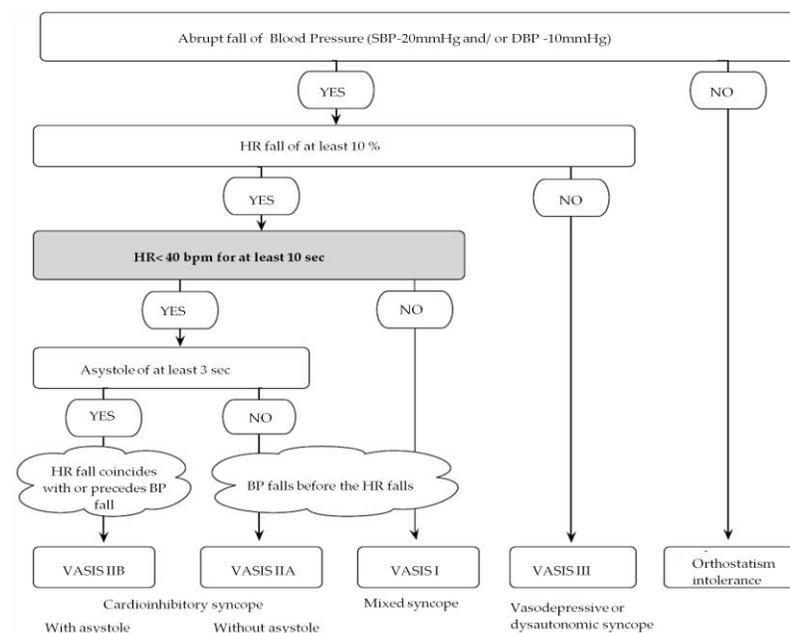


FIG. 2.3- Another classification of reflex syncope based on (Brignole et al., 2000).

Figure 2.4 illustrates the evolution of heart rate and blood pressure in these three types of reflex syncope.

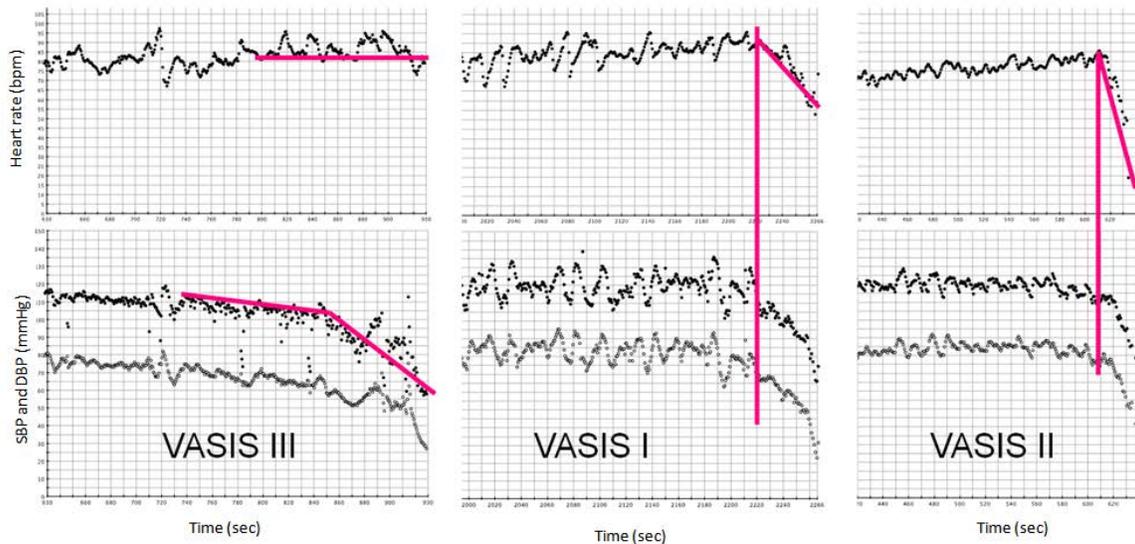


FIG. 2.4 - Example of heart rate and blood pressure evolution in the different types of reflex syncope (DBP, diastolic blood pressure in gray, SBP: systolic blood pressure in black) (Sutton et al 1992).

2.5. Head-Up tilt test

2.5.1. Introduction

The tilt test is a medical procedure commonly used in the diagnostic of syncope (Figure 2.5). It is particularly useful in the diagnostic of vasovagal syncope. This test aims to reproduce the vago-sympathetic disturbances observed during syncope.

In 1986, the tilt test was introduced by Kenny et al. in the clinical evaluation of patients with syncope of unknown origin (Kenny et al., 1986). After that, many protocols have been reported with changes in the initial stabilization phase, the duration of the test, the angle of inclination and different pharmacological prescriptions. It emerges that no standard protocol really exists. We will see in the next chapter the protocol used in this study. Sensitivity and specificity of different protocols are described in detail in several reviews (Brignole et al. 2004) (Benditt et al. 1996).

2.5.2. Methodology

The transition from lying to standing position causes a movement of 500 to 1000 ml of the blood from the thorax to the subdiaphragmatic venous system. When the vertical position is maintained over 10 minutes, the raising of the hydrostatic pressure

leads to a reduction about 15-20% of intravascular volume (700 ml) (Smit et al. 1999). Despite the decrease in venous return and cardiac output, the blood pressure is maintained through compensatory mechanisms.

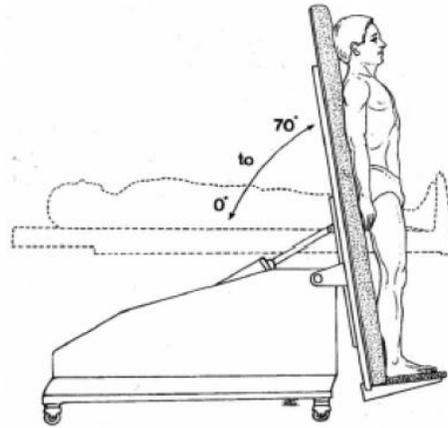


FIG. 2.5 - The Tilt Test (Kapoor, 1990) .

The reflex described by Bezold Jarisch (Mosqueda-Garcia et al., 2000) is at the origin of a bradycardia and a hypotension induced by the tilt test (Figure 2.6). Indeed, tilting the table causes a reduction in venous return and a drop in blood pressure. The compensation mechanisms responsible for increasing the catecholamines in circulating induce vigorous contraction of the left ventricle and over-stimulation of ventricular baroreceptors. This over-stimulation causes inappropriate afferent signals transmitted to the brain stem, and bradycardia and / or hypotension via the efferent pathways (vagal and sympathetic) in subjects susceptible to fainting.

Bezold-Jarisch Reflex

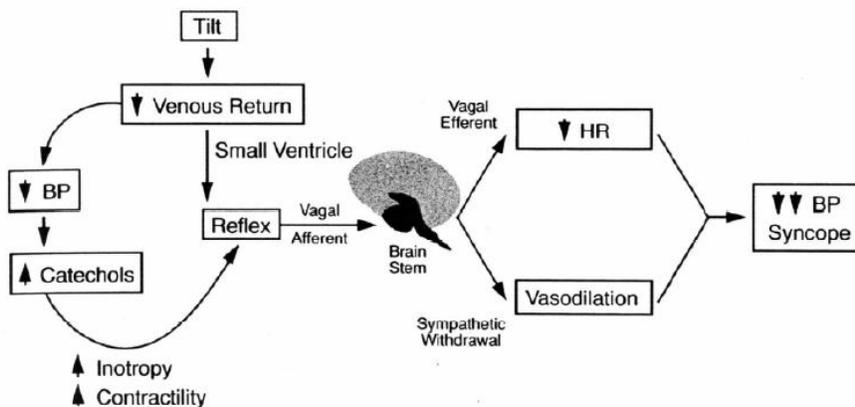


FIG. 2.6 - Reproduction of the Bezold-Jarisch reflex at HUTT according to (Kapoor, 1990).

2.5.3. Complications and contra-indications

The tilt test is not dangerous, it was reported no deaths during the test. However, some rare ventricular arrhythmias have been reported in a pharmacological stimulation (*i.e.* isoproterenol). Minor side effects are common as palpitations and headache. Atrial fibrillation can occur during or after the development of syncope and is usually reversible spontaneously (Castro et al. 2006). Despite the low risk test, its duration (generally between 45 and 60 minutes with the presence of a doctor) may be unsuitable for subjects who are physically weak and do not support the intensification of syncope's symptoms. These problems have motivated the search for methods to predict the outcome of the test at an early stage to minimize the duration of the clinical examination and reduce the economic cost.

2.6. Bibliographic study of syncope prediction

Many methods have been proposed for the early prediction of the outcome of tilt test. Most of them analyze the heart rate variability and blood pressure signal. In 1994, (Madrid et al., 1994) compared a set of time and frequency parameters extracted from 24 hour Holter monitoring with the tilt test results in 50 subjects with a positive response to tilt test (they developed syncope) and 23 non responders (they did not develop syncope). The percentage number of successive RR-intervals which differ by more than 50 ms (PNN50) was identified as the best statistical predictor of the results of tilt test where the sensitivity (Se) which is the percentage of real positive subjects who are correctly classified in positive group, is equal to 51.8% and the specificity (Sp) that computes the percentage of real negative subjects correctly identified such, is equal to 82.6%, while, spectral parameters have shown low predictive capacity. Then (Lippman et al. 1995) analyzed changes in heart rate during the tilt test with the assumption that any increase in parasympathetic tone in response to orthostatic stress is responsible for developing vasodepressor syncope. This study included 28 patients (11 negatives and 17 positives without isoproterenol infusion and 10 positives under isoproterenol infusion). Conventional statistical analysis of the parameters of heart rate variability has been performed and a sensitivity of 41% and a specificity of 100% have been reported. During the rest period, no statistical differences were found between the two groups, while during the tilt, heart rate variability reflected by the square root of the mean squared differences between adjacent intervals period (rMSSD) decreased significantly in comparison with the resting period only in negative subjects, whereas the average RR-interval decreased in a similar manner in both groups.

The increasing in heart rate during the first minutes of the test (after the tilt position) was suggested by (Mallat et al., 1997) as a predictive parameter of positive tilt test. The experiment was divided into two studies:

- Study 1 (without the intervention of a drug) on 110 patients.
- Study 2 (with isoproterenol infusion) on another 109 patients.

The results analyzing heart rate gave a sensitivity of 88.8% and a specificity of 100% for patients in study 1 where the HR has increased less than 18 bpm during the first 6 minutes of the tilted position at 60°. According to this study, (Sumiyoshi et al., 2000) decided to identify whether the increase in HR can predict positive response to tilt test as proposed by (Mallat et al., 1997). They used the same methodology in a new protocol (tilt position at 80° for 30 minutes and without pharmacological treatment) in 115 patients of which 29 were positive (16 of them showed syncope during the first 15 minutes). In this study, the analysis of HR leads to sensitivity=76% and specificity= 62%. However, according to (Turk et al., 2010), increased heart rate during the first 10 minutes of the tilt test has a limited capacity for prediction. A retrospective analysis (Movahed Hassapoyannes and 2001) showed that an increase of HR less than or equal to 18 beats per minute (bpm) for 20 seconds during the first 6 minutes of the tilt test predicts a negative response to tilt test. In 110 patients the results reach a sensitivity =75% and a specificity= 65%.

The use of blood pressure in predicting the results of tilt test is introduced by (Bellard et al., 2001). Its objective was to determine the powerful of early changes in blood pressure in combination with HR. They showed a decreasing in systolic blood pressure and in differential arterial pressure in positive subjects. This decrease in systolic pressure ($> - 6\text{mmHg}$) or in differential arterial pressure in positive subjects at the first 5 minutes of the tilt, showed a sensitivity of 53% and 66% and a specificity of 68% and 67% respectively. Despite the difference between the two groups during the first 5 minutes, none of the heart rate or the blood pressure values can be used to predict the positive results of tilt test (positive predictive value⁵ = 37% and 41% respectively). (Pitzalis et al., 2002) studied the prediction of positive responses to the tilt test on a database of 318 patients with unexplained syncope, by measuring blood pressure before and after the tilt test. A reduction in blood pressure during the first 15 minutes after tilt was observed in positive subjects (sensitivity=58%, specificity=93%).

(Virag et al., 2007) published one of the biggest studies in terms of number of patients. Heart rate and blood pressure were monitored continuously during the tilt test

⁵ Positive predictive Values = number of true positives / (number of true positives + number of false positives).

and for 180 seconds before the test on a database of 1155 patients (759 positive and 396 negative). The signals were processed and combined to develop a model with incremental risk that compares to a threshold for a predetermined vasovagal syncope. The weights assigned to various signals for contribution in predicting syncope were more relevant for blood pressure than heart rate, resulting a sensitivity=95% and a specificity=93%. However, 51% of predictions were in the final minute before syncope, which is a large limit in terms of prediction.

The use of transthoracic impedance (TI) has been introduced by (Schang et al., 2007). They compared a set of parameters measured at rest using support vector machine classifier in 128 patients, 65 had a positive response. These parameters produced 94% of sensitivity and 79% of specificity, without pharmacological intervention. Other studies have analyzed the vasovagal syncope using the aforementioned signals (Stewart et al. 1996) (Kochiadakis et al. 1997) (Liu et al. 2000) (Suzuki et al. 2003), but without the study of the early prediction of syncope.

2.7. Conclusion

We observed that the management of syncope is not an easy task; it is a common obstacle to the health care systems (Chen et al., 2008). In this chapter, we introduced and explained the main ideas behind syncope based on one of the most known clinical guidelines (Moya et al., 2009) who have helped to improve the syncope's management. In the past decade, intensive efforts were made to early predict syncope and to better understand its evaluation process; the last part of this chapter reported the most important outcomes in this field. However, some recent studies have highlighted the difficulties faced when the results of early prediction must be reproduced in various databases, and (Gimeno-Blanes et al., 2011) have concluded that the use of a system for predicting the outcome of the tilt test with existing knowledge should be viewed with caution and a greater effort to better understand the mechanisms of vasovagal syncope. However, we noticed that only few studies have evaluated the non-linear characteristics and the dynamical evolution of time series extracted from data recorded during HUTT. The next chapter will describe the data mining framework that we proposed to investigate this unexplored part.

References

- Bellard, E., Fortrat, J.O., Vielle, B., Dupuis, J.M., Victor, J., Lefthériotis, G., 2001. Early predictive indexes of head-up tilt table testing outcomes utilizing heart rate and arterial pressure changes. *Am. J. Cardiol.* 88, 903–906, A8.
- Benditt, D.G., 2006. Syncope Management Guidelines at work: first steps towards assessing clinical utility. *Eur. Heart J.* 27, 7–9. doi:10.1093/eurheartj/ehi626
- Benditt, D.G., Ferguson, D.W., Grubb, B.P., Kapoor, W.N., Kugler, J., Lerman, B.B., Maloney, J.D., Ravielle, A., Ross, B., Sutton, R., Wolk, M.J., Wood, D.L., 1996. Tilt table testing for assessing syncope. *J. Am. Coll. Cardiol.* 28, 263–275. doi:10.1016/0735-1097(96)00236-7
- Benditt, D.G., van Dijk, J.G., Sutton, R., Wieling, W., Lin, J.C., Sakaguchi, S., Lu, F., 2004. Syncope. *Curr. Probl. Cardiol.* 29, 152–229. doi:10.1016/j.cpcardiol.2003.12.002
- Bloomfield, D.M., Sheldon, R., Grubb, B.P., Calkins, H., Sutton, R., 1999. Putting it together: a new treatment algorithm for vasovagal syncope and related disorders. *Am. J. Cardiol.* 84, 33Q–39Q.
- Brignole, M., Alboni, P., Benditt, D.G., Bergfeldt, L., Blanc, J.-J., Thomsen, P.E.B., Dijk, J.G. van, Fitzpatrick, A., Hohnloser, S., Janousek, J., Kapoor, W., Kenny, R.A., Kulakowski, P., Masotti, G., Moya, A., Raviele, A., Sutton, R., Theodorakis, G., Ungar, A., Wieling, W., 2004. Guidelines on management (diagnosis and treatment) of syncope – Update 2004 The task force on Syncope, European Society of Cardiology. *Eur. Heart J.* 25, 2054–2072. doi:10.1016/j.ehj.2004.09.004
- Brignole, M., Menozzi, C., Del Rosso, A., Costa, S., Gaggioli, G., Bottoni, N., Bartoli, P., Sutton, R., 2000. New classification of haemodynamics of vasovagal syncope: beyond the VASIS classification. Analysis of the pre-syncopal phase of the tilt test without and with nitroglycerin challenge. *Vasovagal Syncope International Study. Eur. Eur. Pacing Arrhythm. Card. Electrophysiol. J. Work. Groups Card. Pacing Arrhythm. Card. Cell. Electrophysiol. Eur. Soc. Cardiol.* 2, 66–76.
- Castro, R.R.T. de, Mesquita, E.T., Nobrega, A.C.L. da, 2006. Parasympathetic-mediated atrial fibrillation during tilt test associated with increased baroreflex sensitivity. *Europace* 8, 349–351. doi:10.1093/europace/eul024
- Chen, L.Y., Benditt, D.G., Shen, W.-K., 2008. Management of Syncope in Adults: An Update. *Mayo Clin. Proc.* 83, 1280–1293. doi:10.4065/83.11.1280
- Gimeno-Blanes, F.J., Rojo-Álvarez, J.L., Caamaño, A.J., Flores-Yepes, J.A., García-Alberola, A., 2011. On the feasibility of tilt test outcome early prediction using ECG and pressure parameters. *EURASIP J. Adv. Signal Process.* 2011, 33. doi:10.1186/1687-6180-2011-33

- J Gert van Dijk, R.S., 2008. Is there any point to vasovagal syncope? *Clin. Auton. Res. Off. J. Clin. Auton. Res. Soc.* 18, 167–9. doi:10.1007/s10286-008-0484-x
- Kabat H, Anderson JP, 1943. Acute arrest of cerebral circulation in man: Lieutenant ralph rossen (mc), u.s.n.r. *Arch. Neurol. Psychiatry* 50, 510–528. doi:10.1001/archneurpsyc.1943.02290230022002
- Kapoor, W.N., 1987. Evaluation of syncope in the elderly. *J. Am. Geriatr. Soc.* 35, 826–828.
- Kapoor, W.N., 1990. Evaluation and outcome of patients with syncope. *Medicine (Baltimore)* 69, 160–175.
- Kapoor, W.N., Peterson, J., Karpf, M., 1986. Defecation syncope. A symptom with multiple etiologies. *Arch. Intern. Med.* 146, 2377–2379.
- Kapoor, W.N., Peterson, J.R., Karpf, M., 1985. Micturition syncope. A reappraisal. *JAMA J. Am. Med. Assoc.* 253, 796–798.
- Kenny, R.A., Ingram, A., Bayliss, J., Sutton, R., 1986. Head-up tilt: a useful test for investigating unexplained syncope. *Lancet* 1, 1352–1355.
- Kochiadakis, G.E., Orfanakis, A., Chrysostomakis, S.I., Manios, E.G., Kounali, D.K., Vardas, P.E., 1997. Autonomic nervous system activity during tilt testing in syncopal patients, estimated by power spectral analysis of heart rate variability. *Pacing Clin. Electrophysiol. PACE* 20, 1332–1341.
- Lippman, N., Stein, K.M., Lerman, B.B., 1995. Failure to decrease parasympathetic tone during upright tilt predicts a positive tilt-table test. *Am. J. Cardiol.* 75, 591–595.
- Liu, J.E., Hahn, R.T., Stein, K.M., Markowitz, S.M., Okin, P.M., Devereux, R.B., Lerman, B.B., 2000. Left ventricular geometry and function preceding neurally mediated syncope. *Circulation* 101, 777–783.
- Madrid, A.H., Moro, C., Marín-Huerta, E., Novo, L., Mestre, J.L., Lage, J., Ricoy, E., 1994. [Usefulness of the RR variability in the diagnosis of neurogenic syncope]. *Rev. Esp. Cardiol.* 47, 536–543.
- Mallat, Z., Vicaut, E., Sangaré, A., Verschueren, J., Fontaine, G., Frank, R., 1997. Prediction of head-up tilt test result by analysis of early heart rate variations. *Circulation* 96, 581–584.
- Martin, G.J., Adams, S.L., Martin, H.G., Mathews, J., Zull, D., Scanlon, P.J., 1984. Prospective evaluation of syncope. *Ann. Emerg. Med.* 13, 499–504.
- Movahed, M.R., Hassapoyannes, C.A., 2001. Prediction of non-occurrence of syncope during a tilt-table test by early heart rate variations. *J. S. C. Med. Assoc.* 1975 97, 207–210.
- Mosqueda-Garcia, R., Furlan, R., Tank, J., Fernandez-Violante, R., 2000. The elusive pathophysiology of neurally mediated syncope. *Circulation* 102, 2898–2906.

- Moya, A. et al, 2009. Guidelines for the diagnosis and management of syncope (version 2009) The Task Force for the Diagnosis and Management of Syncope of the European Society of Cardiology (ESC). *Eur. Heart J.* 30, 2631–2671. doi:10.1093/eurheartj/ehp298
- Pitzalis, M., Massari, F., Guida, P., Iacoviello, M., Mastropasqua, F., Rizzon, B., Forleo, C., Rizzon, P., 2002. Shortened head-up tilting test guided by systolic pressure reductions in neurocardiogenic syncope. *Circulation* 105, 146–148.
- Parry, S.W., Tan, M.P., 2010. An approach to the evaluation and management of syncope in adults. *BMJ* 340, c880.
- Savage, D.D., Corwin, L., McGee, D.L., Kannel, W.B., Wolf, P.A., 1985. Epidemiologic features of isolated syncope: the Framingham Study. *Stroke J. Cereb. Circ.* 16, 626–629.
- Schang, D., Feuilloy, M., Plantier, G., Fortrat, J.-O., Nicolas, P., 2007. Early prediction of unexplained syncope by support vector machines. *Physiol. Meas.* 28, 185. doi:10.1088/0967-3334/28/2/007
- Sheldon, R., Killam, S., 1992. Methodology of isoproterenol-tilt table testing in patients with syncope. *J. Am. Coll. Cardiol.* 19, 773–779. doi:10.1016/0735-1097(92)90517-Q
- Sheldon, R.S., Sheldon, A.G., Connolly, S.J., Morillo, C.A., Klingenheben, T., Krahn, A.D., Koshman, M.-L., Ritchie, D., Investigators of the Syncope Symptom Study and the Prevention of Syncope Trial, 2006. Age of first faint in patients with vasovagal syncope. *J. Cardiovasc. Electrophysiol.* 17, 49–54. doi:10.1111/j.1540-8167.2005.00267.x
- Smit, A.A., Halliwill, J.R., Low, P.A., Wieling, W., 1999. Pathophysiological basis of orthostatic hypotension in autonomic failure. *J. Physiol.* 519 Pt 1, 1–10.
- Soteriades, E.S., Evans, J.C., Larson, M.G., Chen, M.H., Chen, L., Benjamin, E.J., Levy, D., 2002. Incidence and Prognosis of Syncope. *N. Engl. J. Med.* 347, 878–885. doi:10.1056/NEJMoa012407
- Stewart, J.M., Erb, M., Sorbera, C., 1996. Heart rate variability and the outcome of head-up tilt in syncopal children. *Pediatr. Res.* 40, 702–709. doi:10.1203/00006450-199611000-00009
- Sumiyoshi, M., Nakata, Y., Mineda, Y., Tokano, T., Yasuda, M., Nakazato, Y., Yamaguchi, H., 2000. Does an early increase in heart rate during tilting predict the results of passive tilt testing? *Pacing Clin. Electrophysiol.* PACE 23, 2046–2051.
- Sutton R, Petersen M, Brignole M, Raviele A and Menozzi C 1992 Proposed classification for tilt induced vasovagal syncBellard, E., Fortrat, J.O., Vielle, B., Dupuis, J.M., Victor, J., Lefthériotis, G., 2001. Early predictive indexes of head-up

- tilt table testing outcomes utilizing heart rate and arterial pressure changes. *Am. J. Cardiol.* 88, 903–906, A8.
- Suzuki, M., Hori, S., Nakamura, I., Nagata, S., Tomita, Y., Aikawa, N., 2003. Role of vagal control in vasovagal syncope. *Pacing Clin. Electrophysiol. PACE* 26, 571–578.
- Tea, S.H., Mansourati, J., L'Heveder, G., Mabin, D., Blanc, J.J., 1996. New insights into the pathophysiology of carotid sinus syndrome. *Circulation* 93, 1411–1416.
- Turk, U., Alioglu, E., Kirilmaz, B., Duygu, H., Tuzun, N., Tengiz, I., Zoghi, M., Ercan, E., 2010. Prediction of head-up tilt test result: is it possible? *Pacing Clin. Electrophysiol. PACE* 33, 153–158. doi:10.1111/j.1540-8159.2009.02605.x
- Virag, N., Sutton, R., Vetter, R., Markowitz, T., Erickson, M., 2007. Prediction of vasovagal syncope from heart rate and blood pressure trend and variability: experience in 1,155 patients. *Heart Rhythm Off. J. Heart Rhythm Soc.* 4, 1375–1382. doi:10.1016/j.hrthm.2007.07.018

Chapter 3

Description of the data mining framework applied to early predict syncope

The previous chapter focused on the problem of syncope and its diagnosis procedure: the head up tilt test. We have emphasized the importance of using cardiovascular signals, especially the electrocardiogram, as a non-invasive tool to quantitatively analyze the outcome of HUTT and to evaluate the activity of the autonomic nervous system (ANS), in order to early predict syncope. From the literature review presented in chapter 2, on the different methods used to predict syncope, two particularly important points may be highlighted:

- A significant amount of information can be extracted from cardiovascular signals and this information may be potentially useful for providing insights on the diagnosis and prediction of syncope. It can also contribute to provide new knowledge.
- The non-linear characterization and the dynamic evolution of some phenomena have not been sufficiently exploited in this field.

In order to overcome the limitations of previous methods, we propose to analyze the outcome of HUTT through a general data-mining framework, as depicted in figure 3.1.

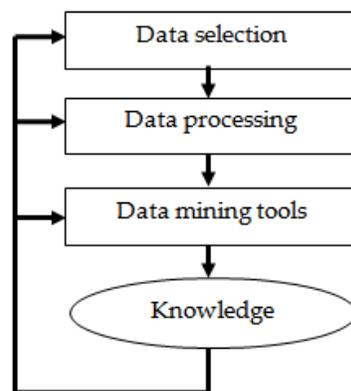


FIG. 3.1 - Data mining framework applied in this manuscript.

The subsequent paragraphs describe how we implement these blocks in this work and the main contributions in terms of engineering sciences and medical knowledge. In the next two sections, we present the database used in this study and the processing algorithms applied to prepare the data for the analysis. Finally, we detail the different feature selection algorithms and classification methods used in this work.

3.1. Data selection

The data are therefore ECGs and other cardiovascular signals such as arterial blood pressure, recorded during head up tilt test from patients who had positive response to the test (they have developed syncope) and others with negative response. Note that this step is already significant. Making a selection of the data reduces the amount of information to be processed, e.g. in case of prediction of syncope, it is more feasible to focus on periods where syncope can be reproduced (during HUTT) rather than long period of Holter recordings. It is also important to make the patient groups as homogeneous as possible, with sufficient size of the database.

3.1.1. Head-Up Tilt Test protocol

Each HUTT was performed between 8:00 to 10:00 AM in a quiet room with dim lights. Subjects were fully informed of the procedure during the inclusion visit 2 to 6 days prior to the HUTT. All participants were asked to avoid physical activity 24 hours and abstain from consuming stimulating beverages (*e.g.* alcohol and coffee) (Karapetian et al., 2012) 12 hours before the test, and to have a light or no breakfast on the day of the test. The HUTT started with 12 minutes of resting in a supine position (6 min rest and 6 min controlled respiration at 0.25Hz), then the table (Sissel, Sautron, France) that has a foot-board support, was tilted at 80° for 45 minutes (Brignole, 2007). No provocative drug was used. The HUTT was considered as positive if the subject developed syncope or intolerable presyncope associated with significant arterial hypotension (decreases of 20 mmHg in systolic blood pressure). In this case, the table was immediately returned to its initial position, and the test was ended. For the positive HUTTs, data were blindly classified independently by three examiners as VASIS I, VASIS II or VASIS III (Sutton et al., 1992)(Brignole et al., 2000). During the HUTT, ECG (standard limb lead configurations), continuous blood pressure using non-invasive finger sensors, cardiac output utilizing impedance technology and autonomic assessment using baroreflex and heart rate variability were monitored with Task Force Monitor® (CN Systems, Graz, Austria). The cuff pressure was also used to provide absolute values of non-invasive blood pressure (figure 3.2).

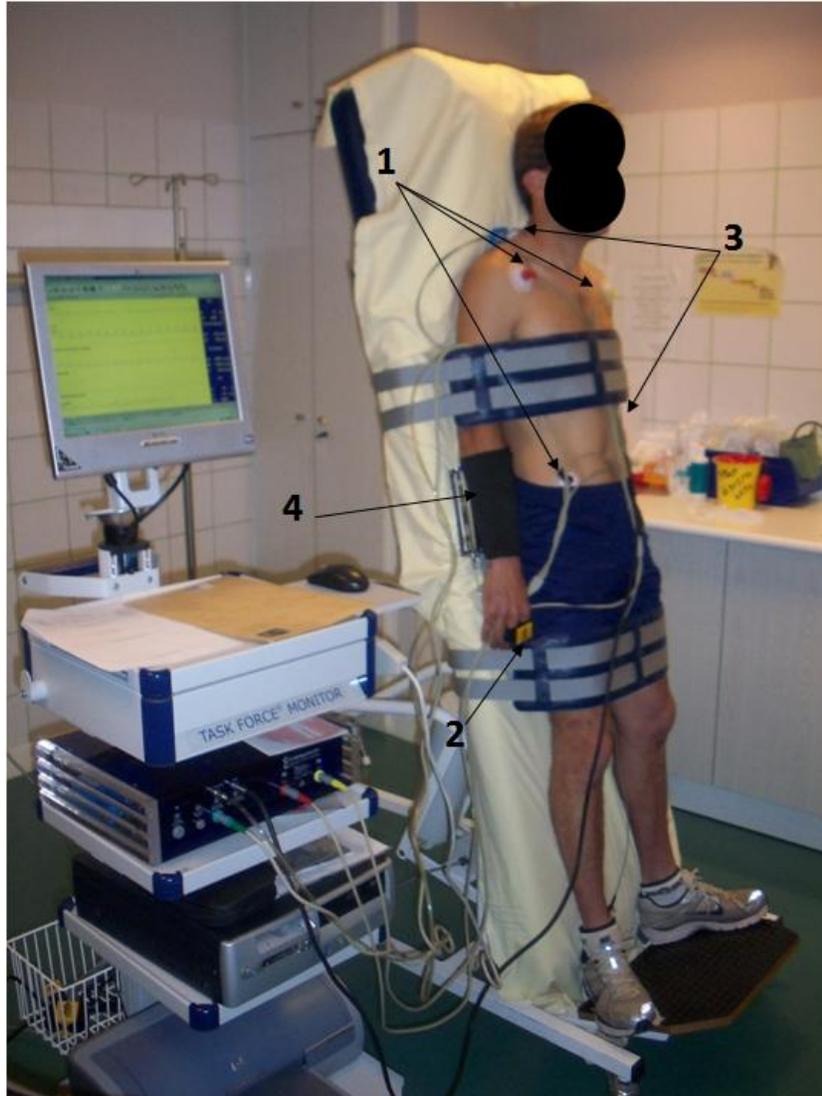


FIG.3.2. Example of head-up tilt test during the inclination phase. The location of different sensors are reported: 1- ECG electrodes, 2- finger sensor to measure blood pressure, 3- impedance cardiography (ICG) electrodes, 4-upper arm cuff

3.1.2. Subjects

Tilt tests were performed on 71 healthy male subjects aged from 18 to 35 years old. They previously underwent an inclusion visit to attest that they were non-smoker, asymptomatic, without classical cardiovascular risk factor (diabetes, dyslipidemia or hypertension), and that they did not take cardioactive medication. From the available 71 subjects, 5 were excluded because of improper ECG signals, 31 subjects were negative to HUTT, whereas 35 were positive: 12 showed a cardio-inhibitory syncope (VASIS II), 7 a vasodepressive syncope (VASIS III), and 16 a mixed syncope (VASIS I). The study was approved by the ethical comity of our University and all participants signed a consent form.

3.2. Data processing

This step consists of removing the undesirable contents from ECG signal (i.e. noise) then extracting useful information. Based on ECG extracted features, a number of blood pressure features are also derived. This information corresponds to the indicators used by cardiologists to characterize the cardiac cycle (including time intervals, wave amplitudes, heart rate, the pulse transit time ...). This extraction step is fundamental because it reduces the size of the data and keeps information interpretable by specialists.

An automatic software for the detection of ECG signals (DELICE), developed and validated by (LTSL, UMR 1099, Rennes) was applied to the raw ECG signal (Dumont et al., 2010). The purpose of this algorithm is to provide a fully automatic segmentation of all heart beats, based on wavelet decomposition, using an evolutionary algorithm for parameter threshold optimization. The data processing includes four main steps that are detailed just after:

1. Preprocessing step: to reduce the noise from raw ECG signals.
2. QRS complex detection: To obtain an accurate and reliable measurement of RR-intervals.
3. Beats averaging : to obtain noise-free beat templates
4. Feature extraction: to obtain pertinent information (Onset, peak and end of the P, QRS and T waves, PR and QT interval...).

3.2.1. Preprocessing signal

The raw ECG was parasitized by various sources of noise: noise related to the patient (movements, breathing) responsible for oscillations of the isoelectric line and the noise associated with the power supply and surrounding devices (50 Hz). The application of filters to the specific raw signal has minimized the noise to get a pre-processed signal.

3.2.2. QRS complex detection

Detection of the QRS complex is obtained by applying the algorithm of Pan & Tompkins (Pan and Tompkins, 1985). The QRS complex corresponds to the highest energy of the ECG signal, which facilitates the development of an automatic detector. The QRS detector includes a high pass filter which attenuates signals of low frequency (P and T waves) and retains the signals of high frequency (which is the QRS complex). Then the detector converts the filtered signal: derivative, squaring and averaging in a time window. Figure 3.2 shows, for one beat, the raw ECG signal and the signal

obtained after transformation. Detecting the time of occurrence of the QRS is defined according to an adaptive threshold. After the adaptive threshold crossed the transformed signal, the algorithm opens a time window in which it searches for the maximum amplitude on the raw ECG signal (Figure 3.3).

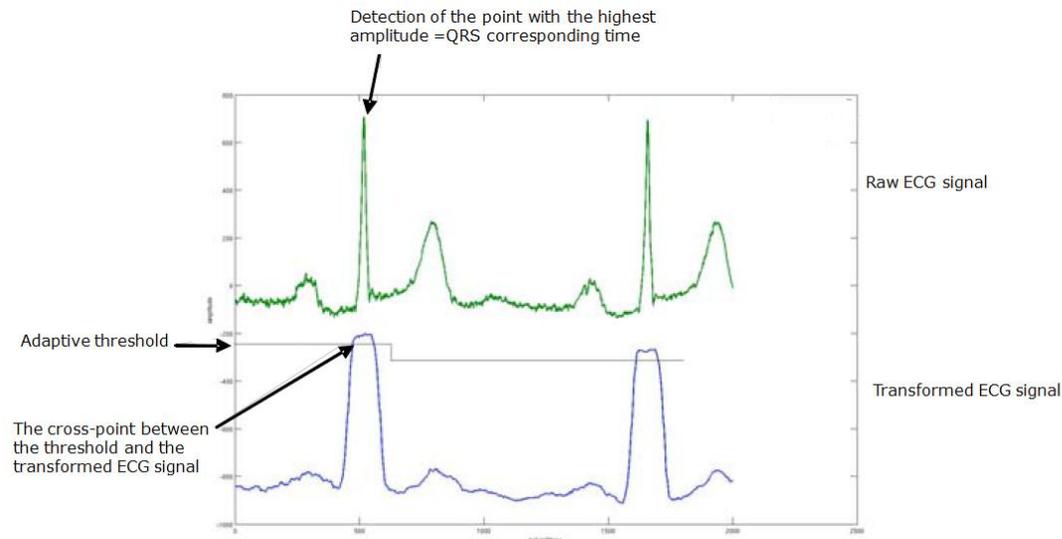


FIG. 3.3 - Example of QRS detection.

The objectives of this phase are in one hand to have a timing reference point for each beat to facilitate subsequent analyzes (averaging and segmentation), and in other hand to obtain an accurate and reliable measurement of all RR-intervals (interval between two successive R waves) of the ECG recording.

3.2.3. Selection and averaging beats

The purpose of this step is to obtain a smoothed beat that will be easier to segment than a noisy beat. First, the algorithm classifies beats according to their morphology then performs an averaging filter on sinus beats.

In the first phase, the algorithm searches for cardiac beats presenting highly-correlated morphologies from the filtered ECG signal. It compares each new detected beat to a set of prototype beats using a correlation indicator (Figure 3.4). Beats presenting a high correlation with a given prototype pattern are grouped into specific populations, so the averaging of the sinus beats is not affected by the presence of ectopic beats. The ECG beats are then aligned based on their QRS complex and averaged within a time window (figure 3.5).

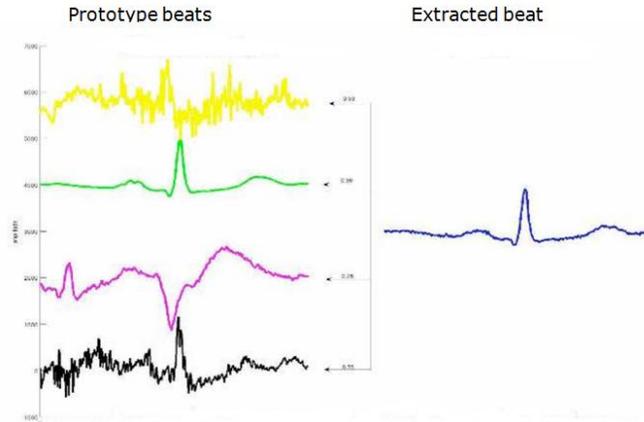


FIG. 3.4 - Left panel: prototype beats corresponding to different morphologies previously detected from the ECG. Right panel: new extracted beat whose morphology is compared with each prototype beat using a correlation function. Example: the extracted beat is weakly correlated with the prototype beats number 1 and 3 (with correlation coefficient of 0.02 and 0.25 respectively) but it is highly correlated with the prototype beat number 2 (with correlation coefficient of 0.9). Therefore it will be included in the prototype group number 2.

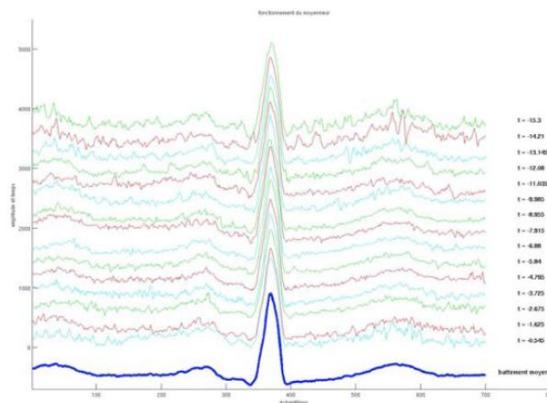


FIG. 3.5 - Alignment and averaging of beats from the same morphology group.

3.2.4. ECG Features extraction

The purpose of this step is to obtain a segmentation of the different waves of an ECG beat by the wavelet transform method based on the work of (Martínez et al., 2004).

Figure 3.6a shows a representation of the typical spectral content of an average ECG beat, identifying the contribution of each ECG wave. We notice that there is a superposition of power spectra of the different waves of the average beat. It is therefore very difficult to segment each wave by a simple filtering. However, as the frequency range of the QRS complex (green curve) is wider than other waves, it is possible to separate a part by filtering.

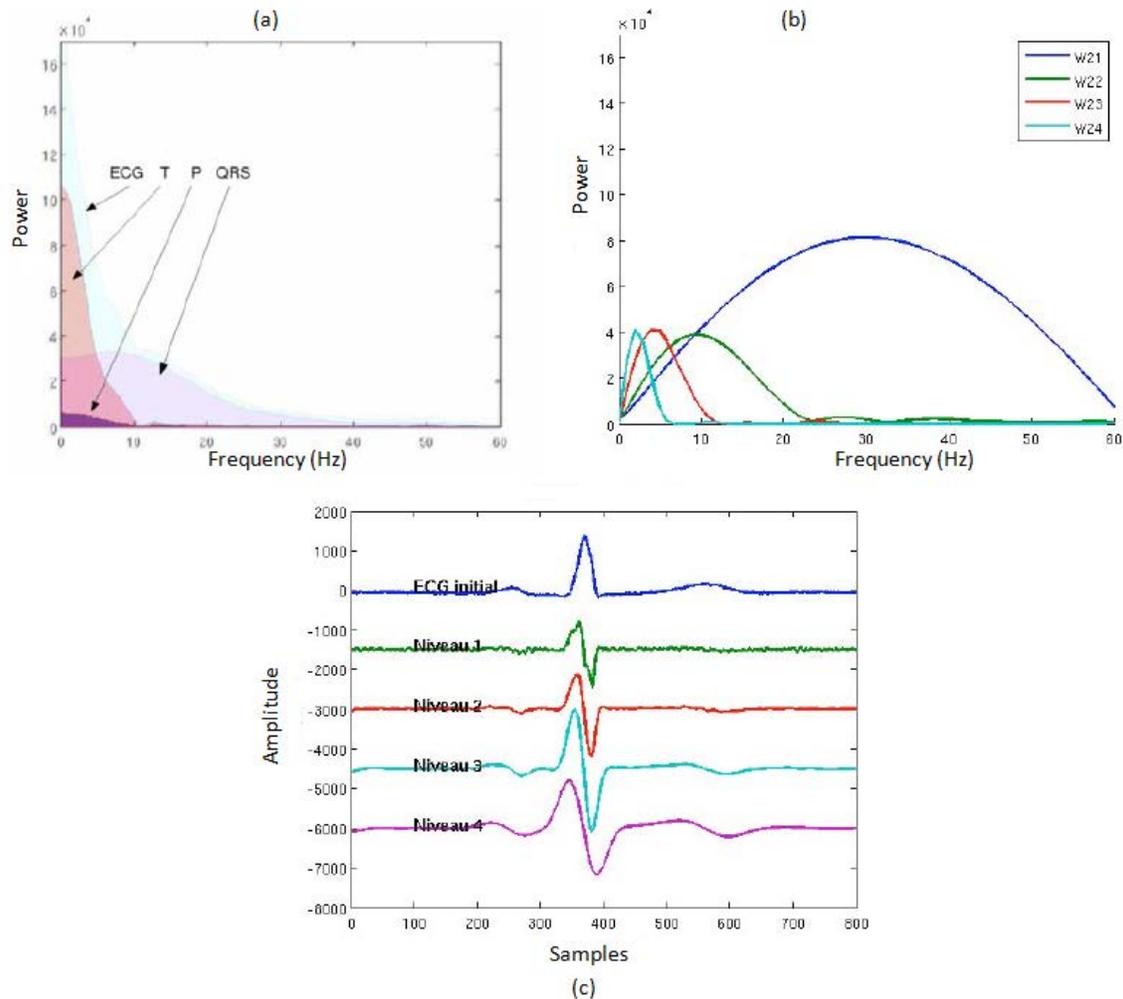


FIG. 3.6 - (a) Spectral representation of a beat and its components: the P, QRS and T wave. (b): presentation of different filters (W21 to W24) applied to the signal to identify a specific frequency range. (c): results from the application of these filters on an initial beat: obtaining a beat decomposed into four specific levels (1- 4).

In wavelet transform method, the ECG beat is decomposed through four different filters (W21-24) each one corresponds to a different spectral band (figure 3.6-b). Two high frequency bands (blue and green) allow the QRS segment and two low frequency bands (red and cyan) permit to segment P and T waves. The initial ECG beat is divided into four levels (1-4) and the algorithm can use the information in different levels of signal to extract the various waves (c). Then, applying different thresholds we can identify the beginning and the end of each wave. Figure 3.7 shows an example of segmentation of the T wave using this method. The beginning of the T wave is detected on the level 4 by crossing the threshold s_1 and the end of the T wave is detected by crossing the threshold s_2 . The peak of the T wave is detected on Level 3 as the moment of zero crossing.

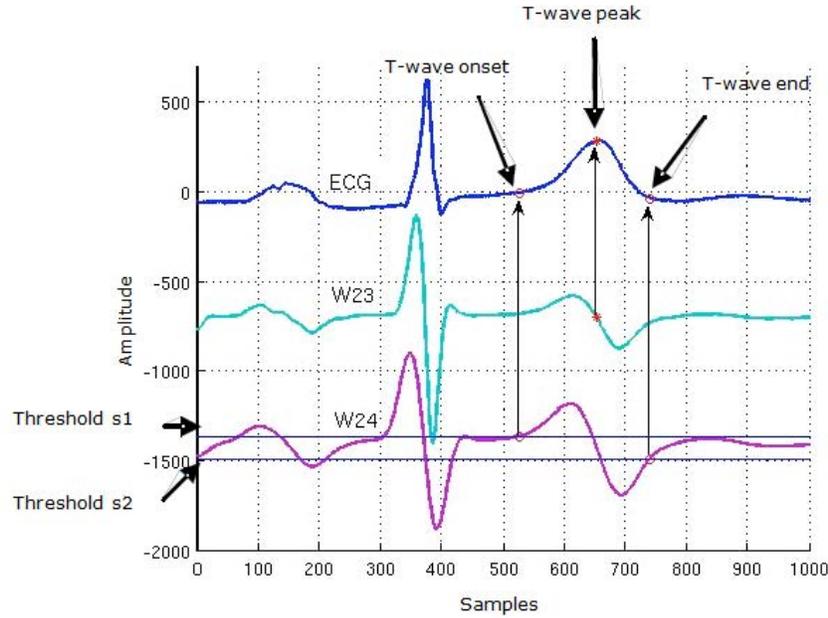


FIG. 3.7 - Example of segmentation of the T wave.

3.2.5. Blood pressure features extraction

It is generally known that blood pressure signal can provide pertinent information about the cardiovascular system, and recently pulse rate variability (PRV) extracted from blood pressure signals has been studied as potential surrogate for heart rate variability (HRV)(Lu et al., 2009)(Charlot et al., 2009). This can be useful in case of the ECG is encompassed with high level of noise or when it is not available (Martin et al., 2003). The principal difference between HRV and PRV is in the time delay, which is caused by the pulse transit time (*PTT*), and is commonly measured as the difference between the appearance of R-peak on the ECG and the corresponding peak value in BP. This value is susceptible to change from one beat to another (Gil et al., 2010).

The systolic (peak) wave is identified from each pulse of the blood pressure through detection methods based on the R-wave of the ECG. First, BP (sampled at 100 Hz) is interpolated using cubic spline, to increase its time resolution and to match that of the ECG signal (sampled at 1000 Hz). Then, each systolic wave in the BP (t_{sj}) was detected as the maximum value of the BP signal within this interval $[t_{Rj}+150\text{ms}, t_{Rj+1}]$ with t_{Rj} is the time occurrence of the R wave (considering that *PTT* is almost 150 ms), see figure 3.8:

$$t_{sj} = \max_{t_{Rj} + 150 \leq t \leq t_{Rj+1}} BP(t) \quad (3.1)$$

After that, the time series of the amplitude of the systolic waves (*AmpS*) was formed:

$$AmpS (j) = BP (t_{p_j}) \quad (3.2)$$

as well as the peak's amplitude variability of the first derivative of systolic blood pressure dP/dt ($dPdt_max$), and the Pulse transit time (PTT), figure 3.8. Beat and pulse detection were then manually verified:

$$PTT (j) = t_{R_j} - t_{s_j} \quad (3.3)$$

The general principle of data transformation makes sense when it is fully included into a data mining framework, which is detailed just after.

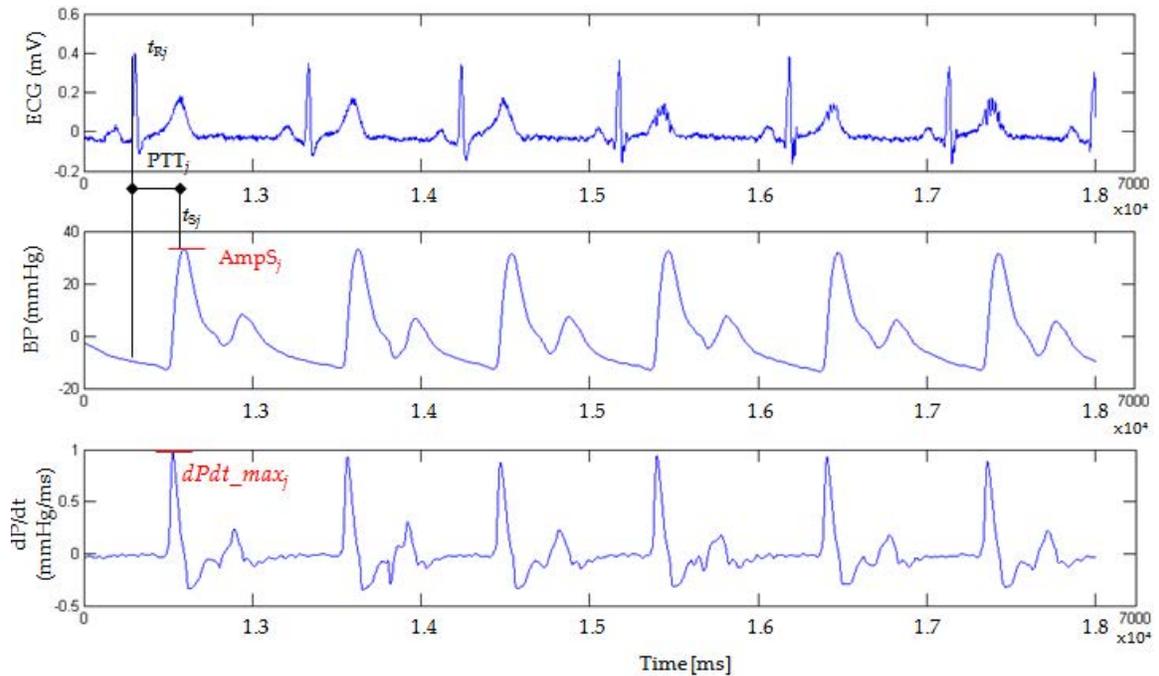


FIG. 3.8 - Definition of the different time series extracted from blood pressure signal

3.3. Data mining tools

This step refers to the extraction of new non-trivial information from large databases. For the particular case of temporal data, the different tasks are generally: (i) prediction, (ii) classification, (iii) clustering, (iv) the search for content, (v) pattern discovery, (vi) visualization and (vii) the detection of anomalies.

The objectives of decision-making in the hospital setting (diagnosis, the choice of treatment) naturally correspond to a problem of classification. However, it is interesting to be able to predict the evolution of the state or the characteristics of a patient, based on current observations or treatment that will be applied to it. For this reason we are

interested in this study to classify the available population into positive and negative groups, which can lead later to an early detection of syncope.

In the past thirty years, the dimensionality of the data participating in machine learning and data mining tasks has increased explosively. Data with extremely high dimensionality (especially with a limited number of patients in the database) has presented serious challenges to existing learning methods. With the presence of a large number of features, a learning model tends to overfit and degenerate its performance. Feature selection is a widely employed technique for overcoming this dimensionality problem. It aims to choose a small subset of the relevant features from the original ones, according to certain relevance evaluation criterion, which usually leads to better learning performance (e.g. higher learning accuracy for classification).

3.3.1. Feature selection

A large number of features can be extracted from cardiovascular signals to participate in the classification task. Sometimes the number of these features can exceed the number of subjects in the database, thus the use of all the descriptors are not feasible. Hence, a selection of the most relevant descriptors that may help in the classification of patients according to their result in the HUTT is needed. Many sophisticated methods can be used to search the most effective features from the available features set. In this manuscript, two well-known and one recently proposed feature selection algorithms have been implemented and tested.

3.3.1.1. The Relief method

Proposed first in (Kenji Kira, 1992), the main idea of relief is to iteratively compute feature weights based on their ability to differentiate between neighboring models. For a given observation O (here is a subject from the database), Relief find its two nearest neighbors: one from the same group (nearest hit) and the other one from a different group (nearest miss). The basic relief algorithm (Kenji Kira, 1992) is explained in figure 3.9 with L the number of features F , m the size of training data and τ is the relevance threshold, which should be between 0 and 1. The 'nu' in the *diff* operation is a normalization unit to normalize the values of *diff* into the interval $[0,1]$. *diff* is used also to calculate the distance between observations to find the nearest neighbors. The sum of differences over all features is taken as the total distance. The relief aims to find the nearest neighbors with respect to important features. Noisy and redundant features may highly affect the selection of nearest neighbors and therefore the selection of features become unreliable. In (Kononenko, 1994) the relief algorithm was extended to increase

the reliability, it differs from the original one by the use of k nearest hits and k nearest misses instead of one nearest hit and one nearest miss and averages their contribution to the weights $W[F]$.

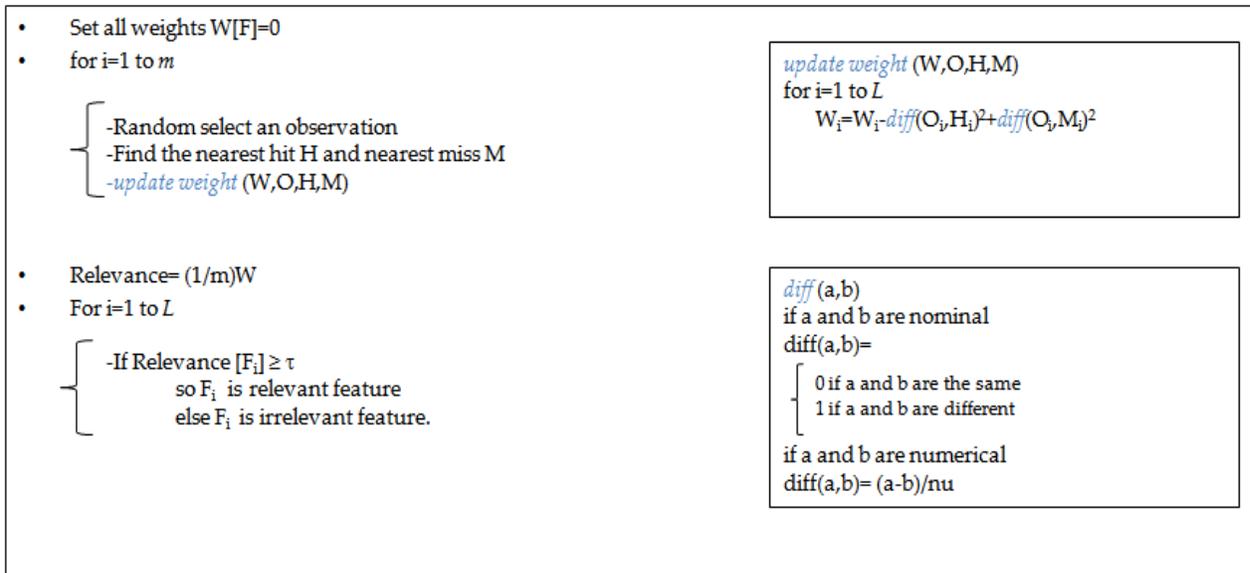


FIG. 3.9 - Relief Algorithm

3.3.1.2. The sequential forward selection (SFS)

Features are added sequentially to a candidate set until the addition of further features does not improve some termination criterion (Devyver and Kittler, 1982). Let $Y=[y_1, y_2, \dots, y_M]$ represents the initial set of features, $X=[x_1, x_2, \dots, x_N]$ ($N < M$) the selected subset with N is the desired number of features, $X \subset Y$ and $J(X)$ the feature selection criterion function for the set X . If we consider that a higher value of J indicates a better feature subset so one possible criterion function is $(1-\mathcal{E})$ where \mathcal{E} denotes the probability of error. This criterion function makes features selection related to the specific classifier used and the training and test data sets. Starting from an empty set, SFS first select a feature x_k with maximum $J(x_k)$ then add sequentially the features x^+ that maximizes $J(x_k, x^+)$ when combined with the features x_k . Figure 3.10-a represents the search space which is drawn like an ellipse to show that there are few features can be selected from the full set. An example of this algorithm is shown in (figure 3.10-b).

3.3.1.3. 'Probe' feature algorithm

It ranks variables using the square of the normalized correlation coefficient between a candidate variable X (dimension N) with x_k , is one of the N observations of X

and the desired output Y with y_k corresponds to the label of patient K (i.e. label 1 if the patient belong to the positive group, label 0, if he belongs to the negative group):

$$r^2 = \frac{(x \cdot y)^2}{(x \cdot x)(y \cdot y)} \quad (3.4)$$

After that, the classification of the variables based on the correlation coefficient was established using the Gram-Schmidt orthogonalization, more details about this method is presented in (Björck, 1994).

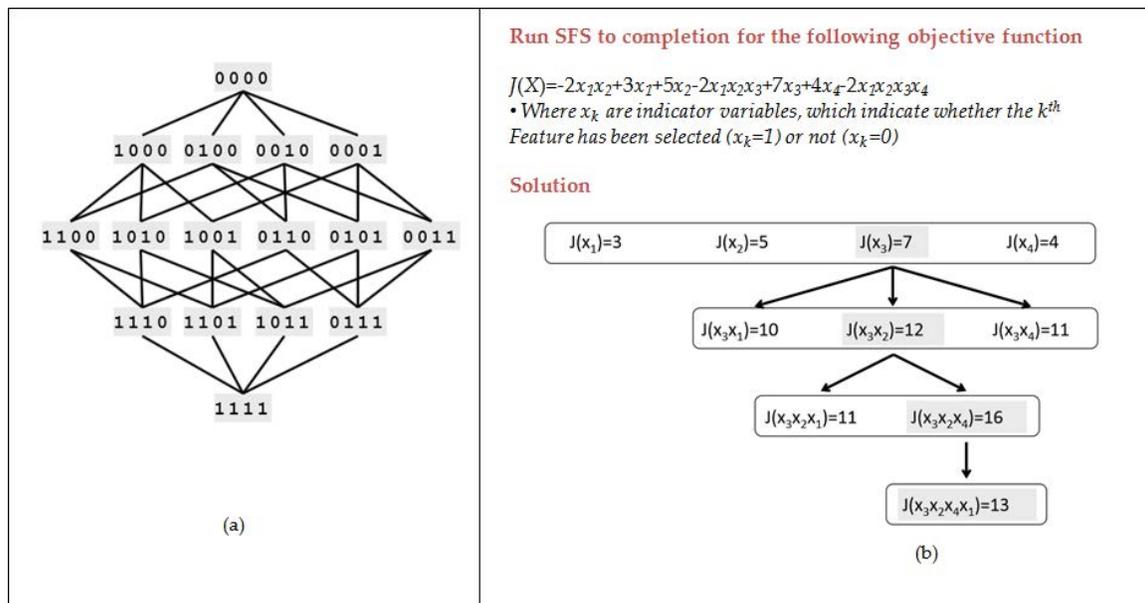


FIG. 3.10 - (a) Representative schema of the elliptic search space of SFS, (b) an example of SFS algorithm.

The 'probe' feature algorithm selects variables, based on the probability of a candidate variable is ranked higher than a random one (Stoppiglia et al., 2003). To estimate this probability, an irrelevant variable or 'Probe' variable is created randomly and added to the set of candidate variables and ranked using correlation coefficient as described above. This procedure is iterated (N times) and the distribution function of the probability that the 'Probe' variable is ranked higher than a candidate one is estimated. Finally the selection of variables is performed by choosing a stopping criterion which is a sufficient value of a risk (α) that a random variable represents the concept more strongly than one of the selected features; an example is represented in (figure 3.11).

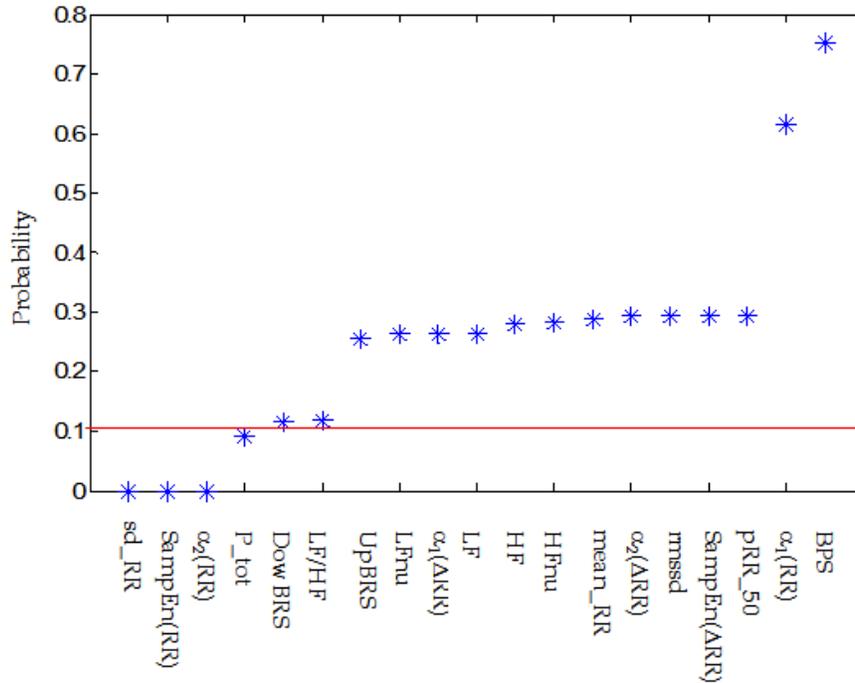


FIG 3.11 - An example of a probability distribution functions that a ‘probe’ feature is better classified than a variable candidate using a set of linear and non linear parameters extracted from the RR-interval time series of subjects during HUTT. Parameters with probability<0.1 (under the red line) are the most pertinent.

3.3.2. Classification

Feature selection algorithms mentioned above look for how sorely a feature is associated with the desired output; they tend to select features that improve the performance of the required classifier. In this manuscript, and in order to separate patients into two classes (negative and positive responses), we evaluate the accuracy of two standard classifiers jointly with feature selection algorithms:

3.3.2.1. K-nearest neighbor (KNN)

It is one of the basic algorithms in machine learning. It assigns a new observation to a class by choosing the majority class among the k nearest examples in the training data.

Algorithm: Given a query instance x_q to be classified:

- Find its k nearest neighbors from the training set measured by a distance function, which can be one of the following equations in Table 3.1.
- Do the majority voting: x_q has to be labeled as the majority label of its neighbors (the class that represents the maximum of the k instances).

TAB. 3.1 - Distance functions

Euclidean	$\sqrt{\sum_{i=1}^k (x_i - y_i)^2}$
Manhattan	$\sum_{i=1}^k x_i - y_i $
Minkowski	$\left(\sum_{i=1}^k (x_i - y_i ^q) \right)^{1/q}$

For our problem, we used the Euclidean distance to measure the distribution between observations, and k is set as the minimum number of nearest neighbors that gives the same ranks of features (with the relief method) starting from $k=\sqrt{N}$ as proposed in (Saravanan, n.d.), where N is the total number of subjects.

3.3.2.2. Support vector machine (SVM)

It can be used to differentiate the data using a decision boundary with maximum margin between the training data and the decision boundary.

Algorithm: given a set of training data of length N (N subjects), $x_i \in \mathfrak{R}^D$ where D is the number of features. For each subject (observation) a label y_i is associated:

$$\begin{cases} y_i = +1 & \text{if } x_i \in \text{class 1 } (C_1) \\ y_i = -1 & \text{if } x_i \in \text{class 2 } (C_2) \end{cases} \quad (3.5)$$

Supposed that the couples (x_i, y_i) , $i=1 \dots N$ are linearly separable in \mathfrak{R}^D with a hyper-plane defined by: $\langle w, x \rangle + b = 0$ with $w \in \mathfrak{R}^D$, $b \in \mathfrak{R}$.

Finding this hyper-plane is equivalent to find the hypothesis function: $f(x) = \langle w, x \rangle + b$ such as:

$$\begin{cases} \langle w, x_i \rangle + b > 0 \Rightarrow y_i = +1 \\ \langle w, x_i \rangle + b < 0 \Rightarrow y_i = -1 \end{cases} \quad (3.6)$$

Vapnik has shown that finding this hyper-plane is equivalent to maximizing the margin (the distance from the hyper-plane to the nearest point x_i) (figure 3.12). This leads to the following optimization problem with constraints:

$$\begin{cases} \min \frac{1}{2} \|w\|^2 \\ \text{with condition} \quad : \quad y_i [\langle w, x_i \rangle + b] \geq 1; \quad i = 1, \dots, N \end{cases} \quad (3.7)$$

Let w^* and b^* be the optimal solutions and α_i are the Lagrange multipliers, the decision function can be written:

$$f(x) = \text{sgn}[\langle w^*, x \rangle + b^*] = \text{sgn} \left[\sum_{i=1}^l y_i \alpha_i^* \langle x_i, x \rangle + b^* \right] \quad (3.8)$$

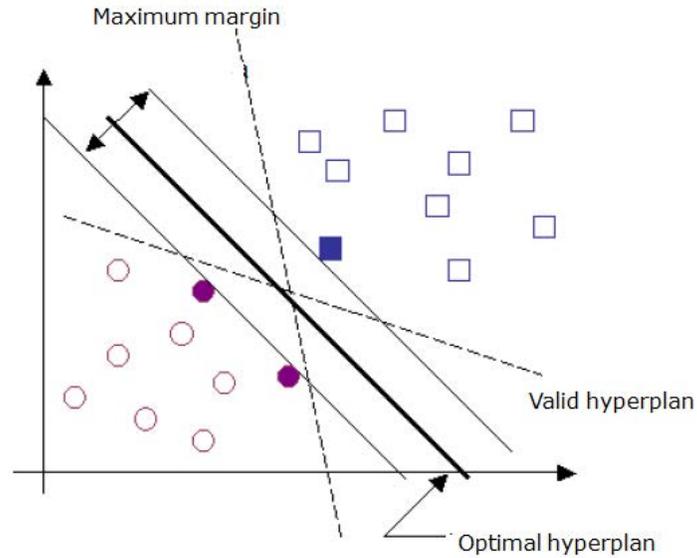


FIG. 3.12 - Classification for optimal hyperplane.

In biomedical research, SVM has proven to be a good model to classify data. However, a classical SVM can separate only data that can be linearly separable. For this reason, SVM can be used with a kernel function to perform a non-linear classification. The idea of kernel support vector machine (KSVM) is to reconsider the non-linear problem in a high-dimensional space, where there likely exists a linear separator. This is the kernel approach, which replaces every dot product by a non-linear kernel function. (Shawe-Taylor and Cristianini 2004). In this case, the decision function becomes:

$$f(x) = \text{sgn} \left[\sum_{i=1}^L y_i \alpha_i^* K(x_i, x) + b^* \right] \quad (3.9)$$

With L is the number of support vectors and K is the kernel function, which can be one of the kernel functions that are already known in the literature (table 3.2).

TAB. 3.2- kernel functions frequently used in the literature

Kernel functions	Equations
Polynomial kernel	$K(u, v) = (u \cdot v + 1)^d$, d is the degree of the polynomial.
Gaussian kernel	$K(u, v) = \exp\left(-\frac{\ u - v\ ^2}{2\sigma^2}\right)$
Exponential kernel distance	$K(u, v) = \exp\left(-\frac{\ u - v\ }{2\sigma^2}\right)$

3.4. Conclusion

This chapter described the main methods applied in this work, presented as a set of steps of a data-mining processing chain. The first two steps (data selection and data processing) are similar for all remaining parts of this manuscript. However, the other tasks differ according to the processing tools used in each chapter. At the end of section 3.2 we obtained multivariate time series consisting of a succession of observations vectors. The objective at this stage is to exploit the non-linear properties and the information contained in the dynamics of these observations. Thus in chapter 4, we compare the evolution of conventional time and frequency parameters of RR-interval time series as well as the baroreflex response in different levels of the tilt test between the positive and the negative test groups. In addition, we explore the changes in dynamic properties of heart rate variability in these two groups, by computing non-linear parameters using sample entropy and detrended fluctuation analysis. In chapter 5, we study the dynamic interaction in time-frequency domain between several cardiovascular parameters that can play important role to improve the understanding of baroreflex control of arterial pressure and to enhance the early prediction of syncope. For this task the smoothed pseudo-Wigner Ville distribution approach is used and two new indexes quantifying the time-frequency relationships between signals are introduced. In chapter 6, we propose a method of analysis of the dynamics of the heart rate variability in the phase space after that we examine the characterization of time series by a set of hidden semi-markov model. The block diagram in figure 3.13 gives an idea of the principle objectives of the remaining chapters.

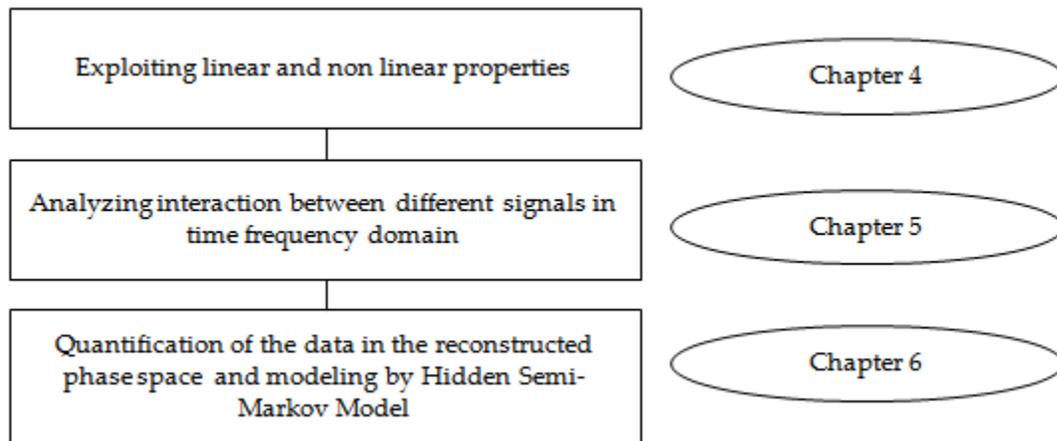


FIG. 3.13 - The main ideas of the remaining chapters.

References

- Björck, Å., 1994. Numerics of Gram-Schmidt orthogonalization. *Linear Algebra Its Appl.* 197–198, 297–316. doi:10.1016/0024-3795(94)90493-6
- Brignole, M., 2007. Diagnosis and treatment of syncope. *Heart* 93, 130–136. doi:10.1136/hrt.2005.080713
- Brignole, M., Menozzi, C., Del Rosso, A., Costa, S., Gaggioli, G., Bottoni, N., Bartoli, P., Sutton, R., 2000. New classification of haemodynamics of vasovagal syncope: beyond the VASIS classification. Analysis of the pre-syncopal phase of the tilt test without and with nitroglycerin challenge. *Vasovagal Syncope International Study. Eur. Pacing Arrhythm. Card. Electrophysiol. J. Work. Groups Card. Pacing Arrhythm. Card. Cell. Electrophysiol. Eur. Soc. Cardiol.* 2, 66–76.
- Charlot, K., Cornolo, J., Brugniaux, J.V., Richalet, J.P., Pichon, A., 2009. Interchangeability between heart rate and photoplethysmography variabilities during sympathetic stimulations. *Physiol. Meas.* 30, 1357–1369. doi:10.1088/0967-3334/30/12/005
- Devyver, P.A., Kittler, J. (1946-), 1982. *Pattern Recognition: A Statistical Approach.* Prentice-Hall.
- Dumont, J., Hernandez, A.I., Carrault, G., 2010. Improving ECG Beats Delineation With an Evolutionary Optimization Process. *IEEE Trans. Biomed. Eng.* 57, 607 –615. doi:10.1109/TBME.2008.2002157
- Gil, E., Bailon, R., Vergara, J.M., Laguna, P., 2010. PTT Variability for Discrimination of Sleep Apnea Related Decreases in the Amplitude Fluctuations of PPG Signal in Children. *IEEE Trans. Biomed. Eng.* 57, 1079–1088. doi:10.1109/TBME.2009.2037734
- Karapetian, G.K., Engels, H.J., Gretebeck, K.A., Gretebeck, R.J., 2012. Effect of caffeine on LT, VT and HRVT. *Int. J. Sports Med.* 33, 507–513. doi:10.1055/s-0032-1301904
- Kenji Kira, L.A.R., 1992. *The Feature Selection Problem: Traditional Methods and a New Algorithm.* 129–134.
- Kononenko, I., 1994. *Estimating Attributes: Analysis and Extensions of RELIEF.* Springer Verlag, pp. 171–182.
- Lu, G., Yang, F., Taylor, J.A., Stein, J.F., 2009. A comparison of photoplethysmography and ECG recording to analyse heart rate variability in healthy subjects. *J. Med. Eng. Technol.* 33, 634–641. doi:10.3109/03091900903150998
- Martin, W., Camenzind, E., Burkhard, P., 2003. ECG artifact due to deep brain stimulation. *The Lancet* 361, 1431. doi:10.1016/S0140-6736(03)13136-4

- Martínez, J.P., Almeida, R., Olmos, S., Rocha, A.P., Laguna, P., 2004. A wavelet-based ECG delineator: evaluation on standard databases. *IEEE Trans. Biomed. Eng.* 51, 570–581. doi:10.1109/TBME.2003.821031
- Pan, J., Tompkins, W.J., 1985. A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng.* 32, 230–236. doi:10.1109/TBME.1985.325532
- Pudil, P., Ferri, F.J., Novovicova, J., Kittler, J., 1994. Floating Search Methods for Feature Selection with Nonmonotonic Criterion Functions, in: *In Proceedings of the Twelveth International Conference on Pattern Recognition, IAPR*. pp. 279–283.
- Saravanan, T., n.d. A Detailed Introduction to K-Nearest Neighbor (KNN) Algorithm. God Your Book Gt.
- Shawe-Taylor, J., Cristianini, N., 2004. *Kernel Methods for Pattern Analysis*. Cambridge University Press.
- Stoppiglia, H., Dreyfus, G., Dubois, R., Oussar, Y., 2003. Ranking a Random Feature for Variable and Feature Selection. *J Mach Learn Res* 3, 1399–1414.
- Sutton, R., Petersen, M., Brignole, M., Raviele, A., Menozzi, C., 1992. Proposed classification for tilt induced vasovagal syncope. *Eur J Cardiac Pacing Electrophysiol* 2, 180–183.

Chapter 4

Kernel based support vector machine for the early detection of syncope during head-up tilt test

As already mentioned in chapter 2 (section 2.6), several methods have been proposed to early predict the results of head-up tilt test. However, few studies have evaluated the non-linear characteristics of HRV during the test (Gimeno-Blanes *et al* 2011). This is due to the fact that most of the non-linear analysis techniques require long-duration time series, greater than those typically recorded in HUTT. Nevertheless, detrended fluctuation analysis (DFA) and sample entropy (SampEn) may overcome this problem, since they may be applied to time series of relatively limited duration. These non-linear indexes have been widely used to the analysis of other cardiac pathologies (Barquero-Perez *et al* 2008, Goya-Esteban *et al* 2008).

This chapter, which is highly inspired from our journal paper published in the *Physiological Measurement* journal (Khodor *et al.*, 2014), aims to compare HRV adaptations during HUTT in subjects with and without syncope, so as to identify parameters that could discriminate these two groups. In the method section, the signal acquisition and processing are presented, then the quantitative analyses of the evaluation of classical (temporal and spectral) as well as non-linear parameters of HRV during HUTT are described, after that a multivariate analysis using kernel-based technique (kernel support vector machine) is proposed in the end of this section to find the optimal hyper-plane that non linearly differentiate positive and negative groups after 15 min of HUTT. The outcome of this chapter is displayed in the results section and is interpreted in the discussion section.

4.1. Methods

The block diagram in figure 4.1 describes our experimental approach. After ECG and blood pressure signals acquisition during the HUTT, a processing step was performed to extract two time series: RR-interval (the time interval between two consecutive R-waves on the ECG) and SS-interval (the time interval between two consecutive systolic waves on blood pressure signals). From these series, we compute linear and non linear parameters, as well as the baroreflex indexes, in the beginning of the tilt test, either on a 5 min or 15 min temporal window. The set of parameters obtained in this manner, are used in the kernel SVM to obtain an output of this model in the form of classification of patients with either a negative or a positive response.

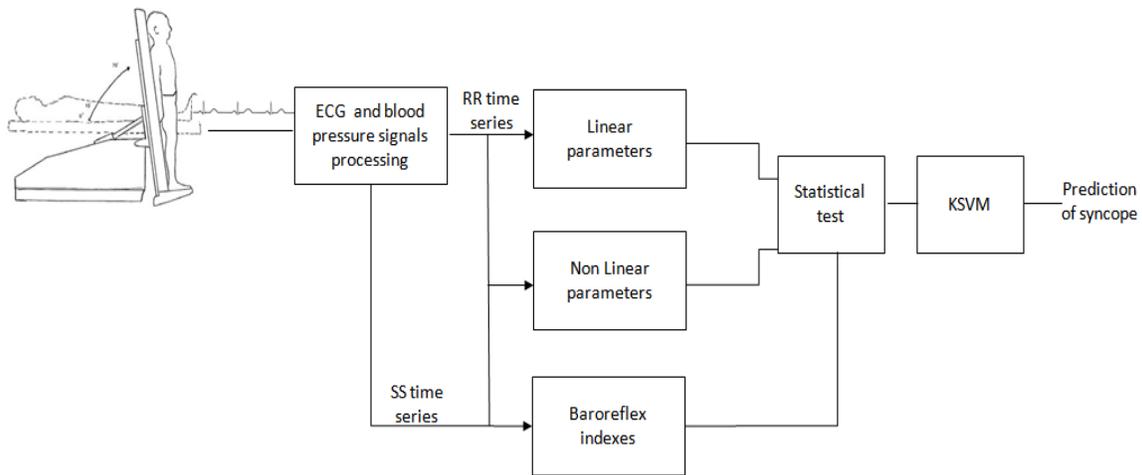


FIG. 4.1 - Block diagram of the proposed approach for syncope prediction.

Head-Up tilt test protocol and subjects description were already reported in chapter 3 (section 3.1). Figure 4.2 shows a characterization plot for heart rate (HR), stroke volume (SV), total peripheral resistances (TPR), mean blood pressure (mBP) and systolic blood pressure (sBP) of a negative and a positive subject during HUTT.

4.1.1. Data processing

4.1.1.1. Preprocessing

One lead ECG and blood pressure signals were acquired at 1000 and 100 Hz sampling frequency respectively. ECG signals were processed with a validated software DELICE (LTSI, UMR 1099, Rennes), which provides a fully automatic segmentation of all heart beats based on wavelet decomposition using an evolutionary algorithm for parameter optimization (Dumont *et al* 2010). This algorithm offers the possibility to consistently extract a large number of ECG indicators (RR-interval, QT-interval ...).

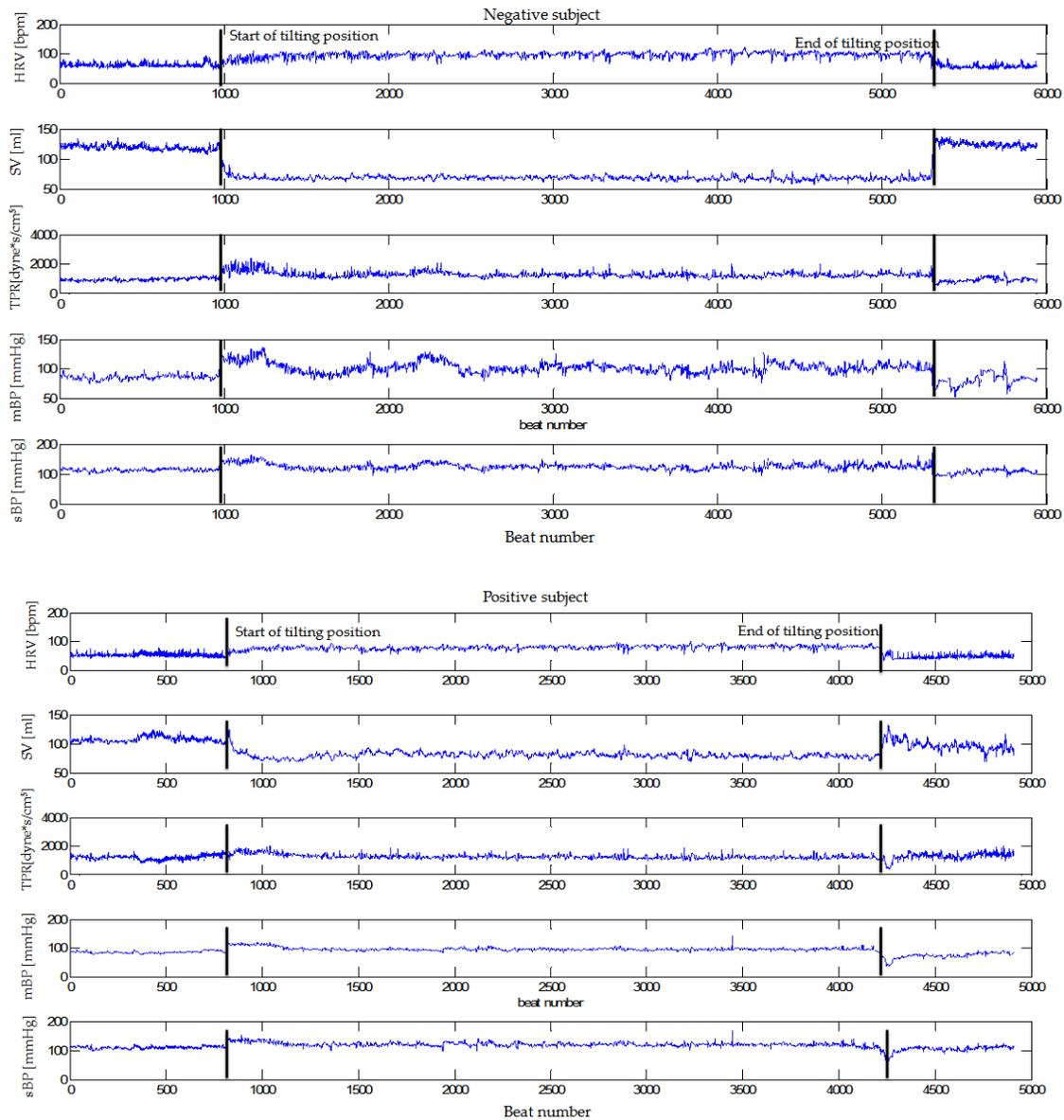


FIG. 4.2 - Characterization plot of different signals measured during HUTT for a negative (upper panel) and a positive (lower panel) subject having syncope of type VASIS II. Abbreviations are as follows: HRV=heart rate variability; SV=stroke volume; TPR=total peripheral resistance; mBP=mean blood pressure; sBP=systolic blood pressure.

It includes also a function for manual correction of possible detection errors. The blood pressure signal $BP(t)$ was interpolated to increase its sampling frequency to 1000 Hz. The temporal positions of the systolic waves tp_j were detected, by finding the maximum of the BP signal on a time support defined between the beginning of the corresponding R-wave (tr_j) plus a fixed duration of 150 ms (which accounts for the electro-mechanical delay and the propagation time until the finger) and the next detected R-wave (tr_{j+1}), (see chapter 3, section 3.2.5) such as:

$$t_{pj} = \max_{t_{Rj} + 150 \leq t \leq t_{Rj+1}} BP(t)$$

The RR-intervals and SS-intervals were extracted from the ECG and pressure signals, on predefined windows, selected from different phases of the HUTT (figure 4.3-4.4):

- During 5 minutes of the supine period with controlled breathing. Only 5 minutes showing similar conditions (controlled breathing) for all subjects were analyzed (Supine5).
- During the first 5 minutes of the tilt test (Early5).
- During the last 5 minutes of the HUTT (Last5):
 - Before syncope occurs; if the patient developed syncope.
 - Before ending the HUTT; if the patient had a negative response.

In order to compute non-linear parameters, we extended the length of the analysis window to 15 minutes; this can be done during tilting position (Early15) and not during supine position because in the latter only 5 minutes of the recording were acquired with controlled breathing.

4.1.1.2. Feature extraction

Temporal indices were represented by the mean and the standard deviation of the RR-interval time series (respectively, mean_RR and std_RR) and the standard deviation of the changes in the heart rate, expressed by the standard deviation of the first derivative of the RR-intervals (rmssd) (Holmestad *et al* 2012). The percentage of pairs of adjacent normal RR-intervals differing by more than 50 ms was also estimated (pRR50). In addition, the noninvasive baroreflex sensitivity (BRS) was estimated as the regression coefficient or slope between the RR-interval time series and the SS-interval time series. An alternative approach called the sequence technique was also used to compute BRS, based on time domain analysis of spontaneous sBP and RR variability, in order to evaluate the sequence in which sBP and RR simultaneously increased (Up-sequence: Up_{BRS}) or decreased (Down-sequence: Down_{BRS}) over three beats (figure 4.5). More details about the sequence technique can be found in (Pitzalis *et al.*, 1998)(Gouveia *et al.*, 2005) (Paola Martínez-García, 2012).

An auto-regressive model (AR) with order 12 (the model order was selected according to Akaike information criterion and the goodness of fit of the AR model was tested using the whiteness test) was applied to compute the RR-interval power spectrum. Spectral indices were calculated by measuring the area under 3 usual frequency bands (Electrophysiology, 1996), low frequency power (LF) with limits: 0.04–0.15 Hz, high frequency power (HF) with limits: 0.15–0.4 Hz and total power spectrum (P_tot) with limits: 0.04–1 Hz.

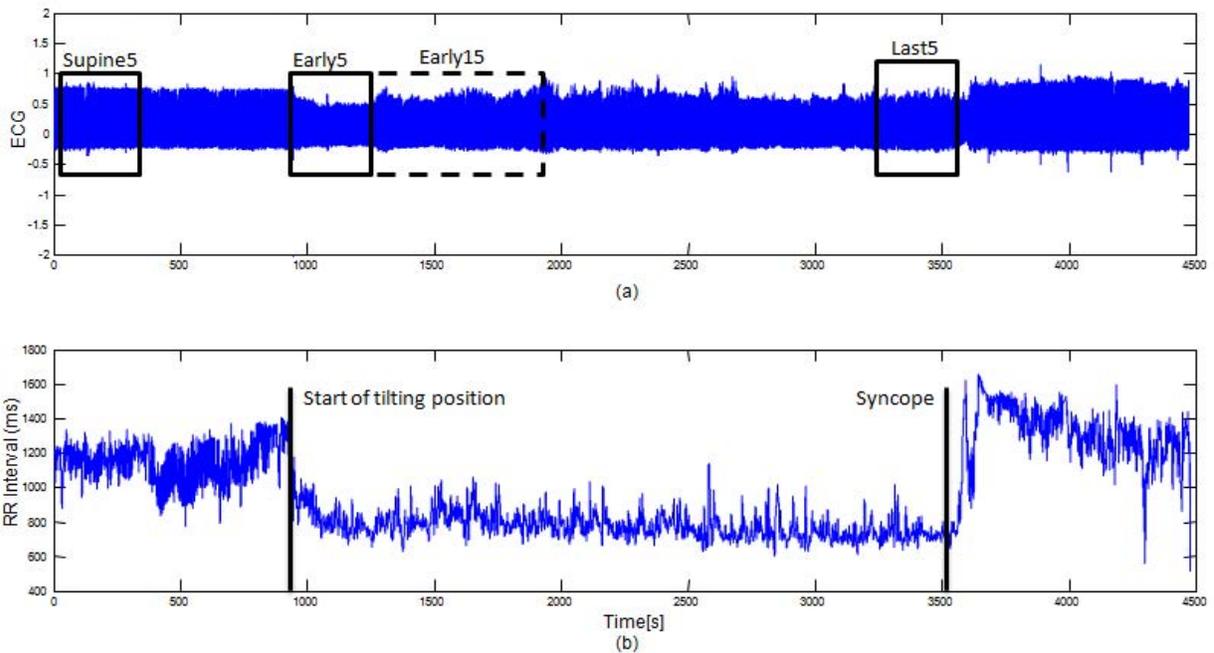


FIG. 4.3 - Example of an ECG recording with the different time intervals used for analysis, for a subject with positive response (a) and the corresponding RR time series (b).The RR-interval time series corresponding to each window are used to extract the features that will be used for classification.

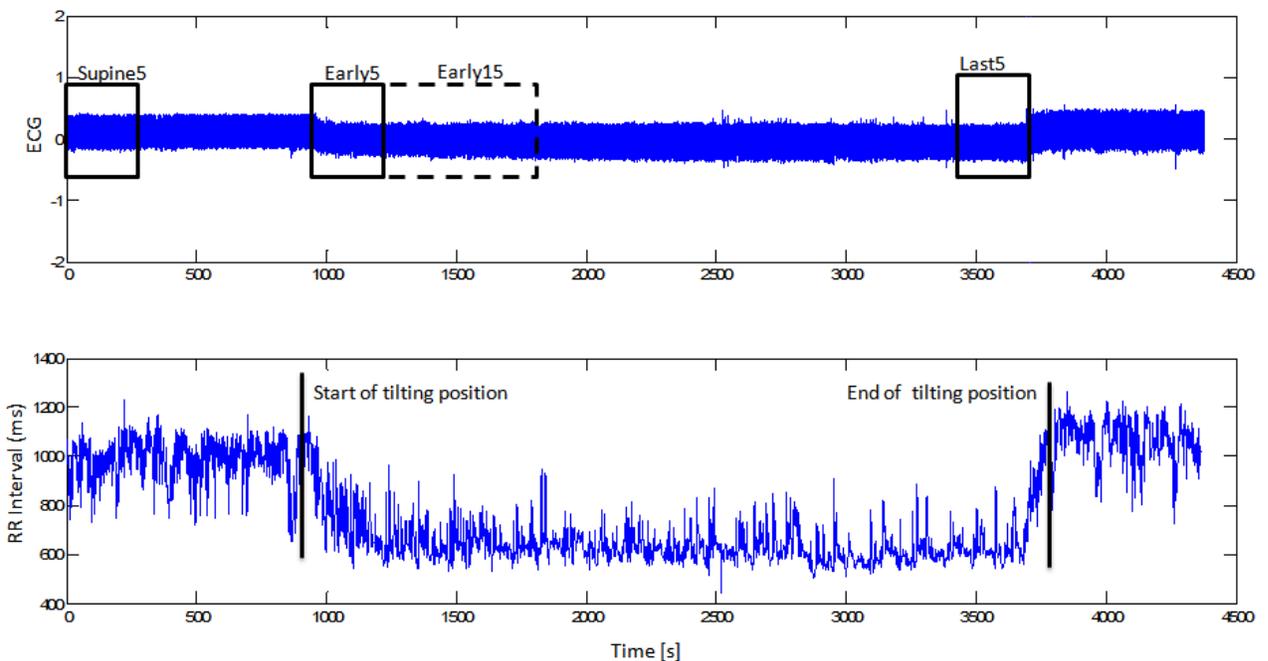


FIG. 4.4 - Example of an ECG recording with the different time intervals used for analysis, for a subject with negative response (a) and the corresponding RR time series (b).The RR-interval time series corresponding to each window are used to extract the features that will be used for classification.

LF and HF indices are expressed in normalized units as LFnu obtained by LF/(HF+LF) and HFnu obtained by HF/(HF+LF), the ratio LF/HF is then computed. To ensure that respiration did not influence LF and HF powers, the respiration frequency was verified in all subjects, and it was maintained between 0.2-0.36 Hz (during the uncontrolled breathing time: Early5, Early15 and Last5) with the 25th and 50th percentile equal to 0.2Hz and 0.23Hz respectively.

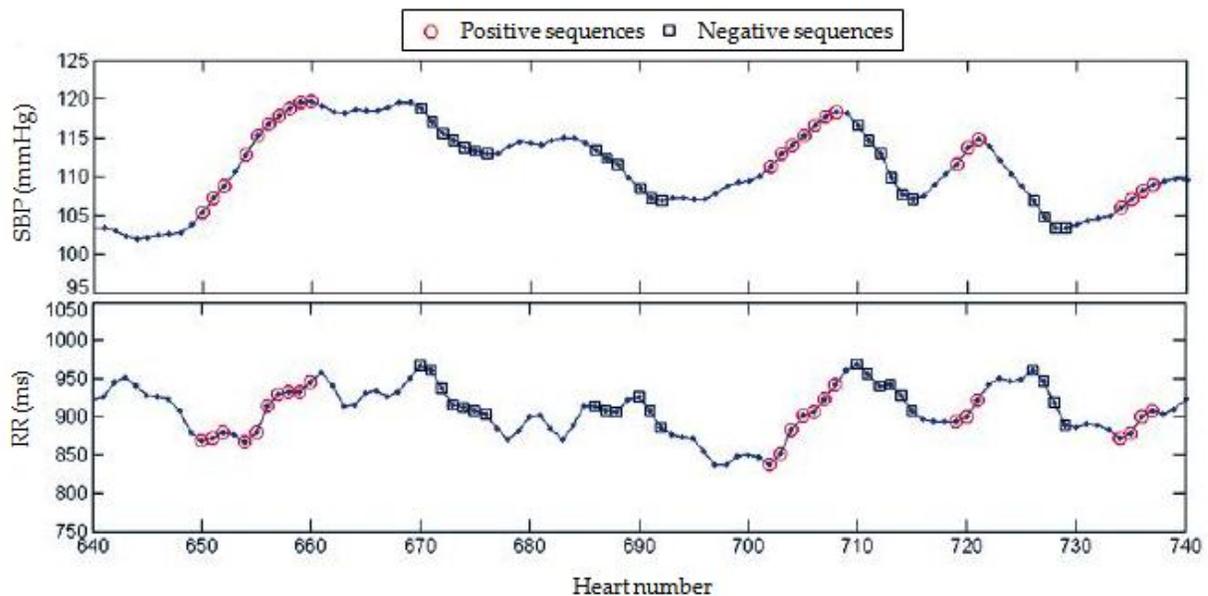


FIG. 4.5 – Example of identification sequences with simultaneous increasing or decreasing in sBP and RR interval (Paola Martínez-García, 2012).

Non-linear parameters, Detrended Fluctuation Analysis (DFA) and Sample Entropy (SampEn) (explained in details in the appendix A) were computed for the RR-interval and the Δ RR-interval (the difference between two successive RR-intervals):

- The DFA algorithm was applied as described by (Goldberger *et al* 2000), quantifying the scaling behavior of RR-interval time series. Referring to previous experiments (Francis *et al* 2002), α_1 was calculated between the range $n=4$ and $n=16$ and α_2 was calculated between the range $n=16$ and $n=64$.
- Sample entropy (SampEn) is a useful tool that illustrates the amount of complexity or irregularity in time series data. It is a modified version of Approximate entropy (ApEn), but unlike ApEn, it overcomes the limitation of biased measures (Richman and Moorman 2000). SampEn is the negative natural logarithm of the conditional probability that subseries of length m having repeated itself within a tolerance r for m points will also repeat itself for $m+1$ points (Lake *et al* 2002). A low value of SampEn arises from a high degree of regularity. We have calculated the SampEn using a maximum epoch length $m=2$ and a tolerance $r=0.2$. These variable ranges

were chosen referring to a previous work (Pincus 1991) that demonstrated good statistical validity for SampEn.

4.1.2. Data Mining

The above-mentioned signal processing tools and data analysis methods were applied to the entire database described in the Methods section. In the data-mining step, an intra-group comparison was first performed, followed by an inter-group comparison on each parameter separately. In order to establish a clinical diagnosis test, an optimal KSVM was researched for a diagnosis purpose to test the following two hypotheses:

- H_0 : no syncope will occur during the tilt test,
- H_1 : the patient will be subject to syncope.

4.1.2.1. Statistical analysis

Statistical calculation was performed using Statistics Toolbox of Matlab 2011a. Intra- and inter-group comparisons among all the variables were performed. First, intra-group comparisons were assessed among the 5 minutes of supine resting (Supine5), the first 5 minutes of tilted position (Early5) and the last 5 minutes before ending the tilt test (Last5). This was done to evaluate whether heart rate indices change during the test, by performing a pair wise comparison between indices during each pair of condition using Wilcoxon signed rank test. Second, for each condition (Supine5, Early5 and Last5) and for the first 15 min of the tilted position (Early15), values were compared between negative and positive test groups (inter-group comparisons) using Mann-Whitney test. A p -value <0.05 was considered as significant.

4.1.2.2. Kernel support vector machine

In order to separate patients into two classes (negative and positive responses), a support vector machine (SVM) was used to differentiate the data using a decision boundary with a maximum margin between the training data and the decision boundary. In biomedical research, SVM has proven to be a good model to classify data. However a classical SVM can separate only data that can be linearly separable. For this reason, SVM was used with a kernel function to perform a non-linear classification. The idea of kernel support vector machine (KSVM) is to reconsider the non-linear problem in a high dimensional space where there likely exists a linear separator. This is the kernel approach, which replaces every dot product by a non-linear kernel function. More details about kernel methods can be found in (Shawe-Taylor and Cristianini 2004).

The effectiveness of KSVM depends on the selection of the kernel and the kernel's parameters. The performance of the KSVM using two of the most commonly used and recommended kernel functions (the radial basis function RBF and the polynomial function) was examined on the computed parameters to compare and find the best discriminator. Overall, RBF kernel provided the best performance. For this reason it was adopted for the rest of the study. After ECG signal acquisition during the tilt test, the signal undergoes a processing step to extract several indicators. From the extracted RR-interval time series we computed linear and non-linear parameters in addition to Baroreflex indexes, which were used in the kernel SVM to obtain an output of this model in the form of classification of patients with either a negative or a positive response.

The database with $K=66$ subjects was divided into a training set and a testing set. To overcome the limited number of subjects and avoid overfitting, a leave-one-out cross-validation (LOOCV) method was applied. One subject from the database is used as validation data and the remaining ($K-1$) subjects are used as training data. This procedure is repeated K -times. Each time a new subject is used to test the KSVM model. Kernel parameters were varied firstly in order to get a minimum error of classification and secondly to obtain an optimal performance with zero error of classification of patients with positive response. The performance of the classification was evaluated in each window separately (Supine5, Early5, Last5 and Early15), trying different combinations of several parameters together, (the choice of parameters was based on exhaustive manner to try all possible combinations) in order to get the best performance in each time interval. After that, the value of kernel's parameter and the combination of parameters that give the best performance in LOOCV in each window were used to perform 2-fold cross validation.

4.1.2.3. Performance evaluation

Performance was evaluated using the classical sensitivity (Se) and specificity (Sp) indexes (chapter 2 section 2.6). Moreover global performance of the accuracy of the classifier is defined as the percentage of correct classifications. Thus, the error (\mathcal{E}) is the percentage of patients that were incorrectly classified by the classifier.

4.2. Results

4.2.1. Intra-group comparison

An intra-group analysis was performed to determine the effect of position change on the evolution of parameters in both negative and positive test groups. To

achieve this goal, Supine5, Early5 and Last5 were used (see figure 4.3). The HUTT procedure has only 5 min of controlled respiration during the supine position, which compelled us to do the intra-group comparison using a 5-minute window length.

In both groups, a significant increase in LFnu and LF/HF was observed from Supine5 to Early5 ($p<0.01$) concomitant with a decrease in HF, HFnu, mean_RR, rmissd, pRR50, BRS, Up_{BRS} and Down_{BRS} ($p<0.01$). These parameters continued to evolve in the same way from Early5 to Last5 excluding rmissd, pRR50, Up_{BRS} and Down_{BRS} ($p>0.05$). LF, P_{tot} and sd_RR did not show any significant change. Figure 4.6 presents the median and the median absolute deviation of the classical and the indices with the baroreflex response during the different term interval for both negative and positive groups.

4.2.2. Inter-group comparison

In Supine5, our results failed to show any significant difference between negative and positive groups in regards to classical (temporal, spectral and BRS) parameters. However, during Early5 and Last5, a significant difference was found between negative and positive test groups in LF ($p<0.05$), where LF was greater in the negative than in the positive test group. Conversely, during the 15 minutes of tilting position, P_{tot}, sd_RR, α_2 (RR), SampEn(RR) and α_2 (Δ RR) ($p<0.01$) were different between groups (Table 4.1). Figure 4.7 illustrates the result of the inter-group comparison tests with negative subjects displaying sd_RR, α_2 (RR) and α_2 (Δ RR) lower than in positive subjects, SampEn(RR) was greater in negative subjects than in positive subjects. Figure 4.8 illustrates two typical examples of DFA parameters in negative and positive patients.

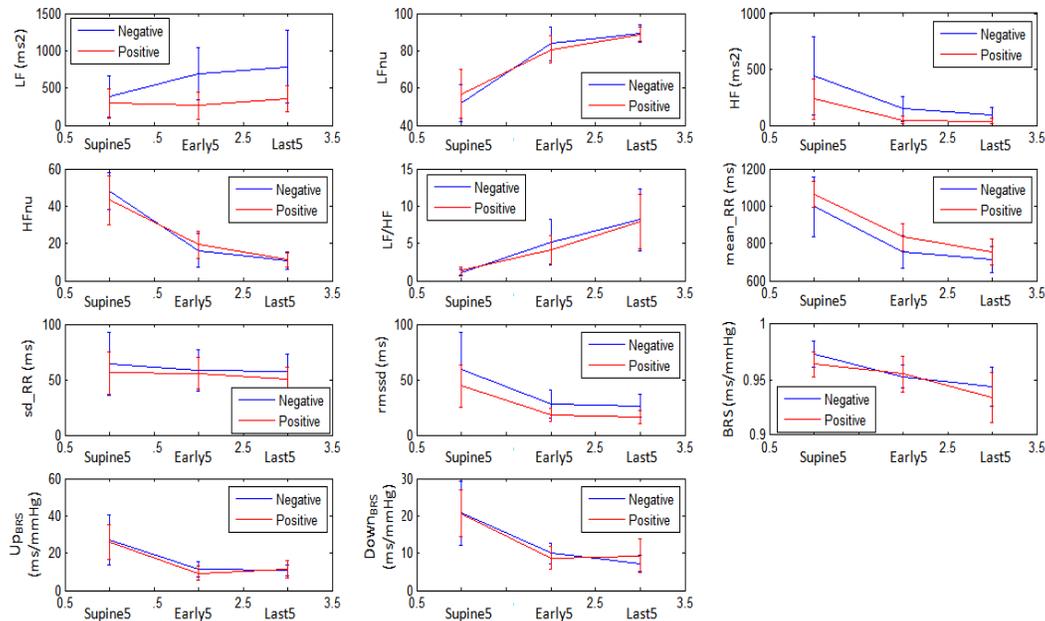


FIG. 4.6 - Median and MAE values of parameters with windows of 5 min.

TAB. 4.1- Median and median absolute deviation values of parameters during the first 15 min of HUTT with the p-value of the Mann-Whitney test. (*) designates significant parameters

Early15	NEG	POS	p-value
LF(ms ²)	702.5±365.3	436.6±304	0.258
LFnu	85.4±6.4	81.7±5.3	0.161
HF(ms ²)	117.2±88.9	120.4±88.4	0.738
HFnu	14.6±6.4	18.3±5.3	0.161
P_tot	1443.3±870.2	1979.2 ±1703.1	0.032 (*)
mean_RR(ms)	759.9±97.2	854.2±85.5	0.052
sd_RR(ms)	56.5±17.6	81.9±42.3	0.006(*)
rmssd(ms)	25.8±10.7	26.7±10.9	0.7972
pRR50	5.8±5.4	7.04±6.4	0.897
SampEn(RR)	0.9±0.2	0.71±0.2	0.0001(*)
SampEn (ΔRR)	1.6±0.2	1.5±0.3	0.217
α ₂ (RR)	0.8±0.14	1.02±0.1	0.0001(*)
α ₁ (RR)	1.2±0.25	1.3±0.2	0.3823
α ₂ (ΔRR)	0.2±0.07	0.3±0.08	0.038(*)
α ₁ (ΔRR)	0.6±0.2	0.6±0.2	0.625
BRS(ms/mmHg)	0.95±0.01	0.94±0.02	0.681
Up _{BRS} (ms/mmHg)	27.22±13.5	25.9±9.3	0.464
Down _{BRS} (ms/mmHg)	20.75±8.6	20.63±6.3	0.227

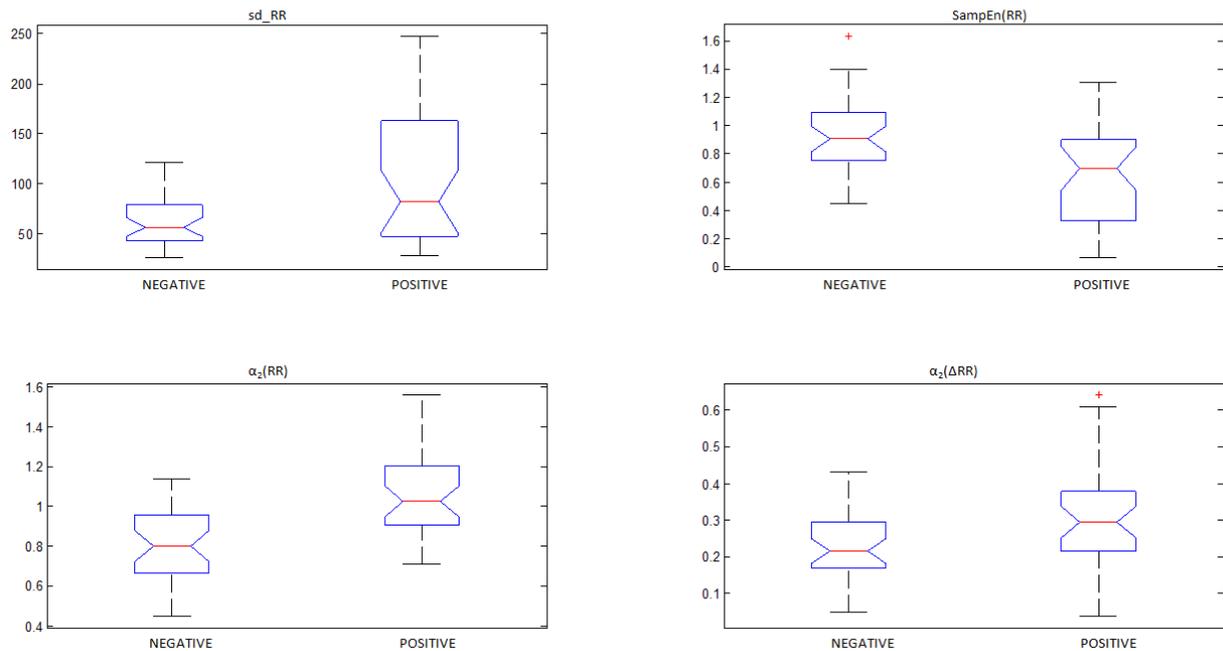


FIG. 4.7 - Boxplot of the significant parameters in Early15.

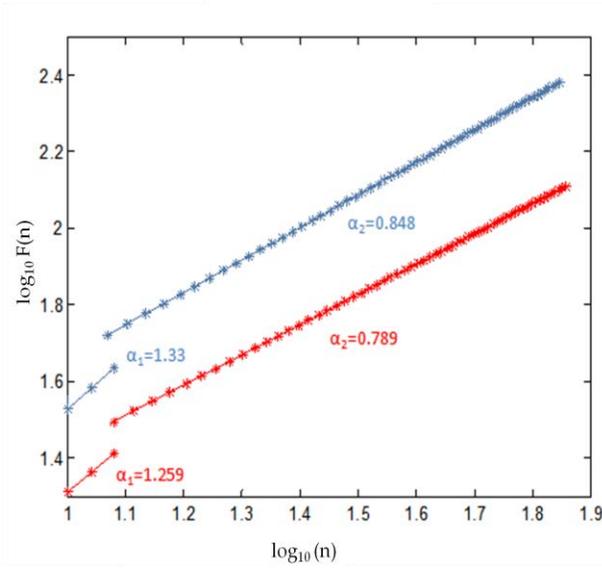


FIG. 4.8 - DFA of RR-interval time series in negative (in red) and positive (in blue) patients examined during the first 15 min of HUTT.

4.2.3. Classification

As mentioned above, KSVM was used to test the performance of the parameters to distinguish between the two groups and to correctly classify all subjects with positive response. Parameters extracted from the RR-interval time series in windows of 5-minutes length (Supine5, Early5 and Last5) and a window of 15-minutes length (Early15) were evaluated to find the window with the best performance of KSVM. We used the radial basis as a kernel function; its parameter (σ) was varied from 0.01 to 4 using 0.01 as step size. With $\sigma > 4$, the error of classification was generally large in all windows, this is why we decided to stop at $\sigma=4$.

- **Performance with minimum error ϵ_{\min} of classification using leave-one-out cross validation (LOOCV).** An exhaustive procedure was used initially in order to get the best performance in each window. Table 4.2 displays the contingency tables when obtaining the minimum error in the different windows (Supine5, Early5, Last5 and Early 15):

- In Supine5, the accuracy of the classifier is 78.7 % (with a sensitivity of 80 % and a specificity of 71%) when we take the four parameters LFnu, HF, LF/HF and mean_RR and apply the radial basis function with $\sigma=0.36$ as kernel parameter (table 4.2-A).
- In Early5, the accuracy of the classifier is lower than in Supine5 (71.2%) using LF, LFnu, LF/HF and mean_RR for $\sigma=0.48$ (with Se=74.3%; Sp=67.7%) (table 4.2-B).

- In Last5, the best performance obtained is 74.3%. We observed a sensitivity Se of 71.4% and specificity Sp of 77.4% when LF, P_{tot} , $mean_RR$ and sd_RR are used with $\sigma=0.32$ (table 4.2-C).
- The best accuracy of the classifier (84.8%) is found in Early15 if P_{tot} , sd_RR , $SampEn(RR)$, and $\alpha_2(RR)$ are used and σ equals to 0.57 with 88.5% of sensitivity and 80.6% of specificity (table 4.2-D).

TAB. 4.2 - Contingency tables in the different windows when a minimum error of classification is obtained in LOOCV with R_NR: the real group of patients with negative response; R_PR : the real group of patients with positive response; K_NR: kernel SVM classification of patients with negative response; K_PR: kernel SVM classification of patients with positive response.

A	Supine 5	
	R_NR	R_PR
K_NR	22	7
K_PR	9	28

B	Early5	
	R_NR	R_PR
K_NR	21	9
K_PR	10	26

C	Last 5	
	R_NR	R_PR
K_NR	24	10
K_PR	7	25

D	Early 15	
	R_NR	R_PR
K_NR	25	4
K_PR	6	31

The individual parameter's evaluation in Early15 indicates that the minimum obtainable error using one parameter is equal to 25 % for $\alpha_2(RR)$, 27.3% for sd_RR and $SampEn(RR)$ and 31.8% for P_{tot} when the radial basis is applied as a kernel function. This indicates that the best performance is obtained using the four parameters simultaneously, which gives independently the minimum error of classification. The results of 2-fold cross validation are shown in table 4.3. In this experiment, 15 negative subjects and 18 positive subjects randomly selected from the whole population were used for the training step. The remaining population was used as testing set. This process was repeated 10 times.

TAB. 4.3- Mean performance using 2-fold cross validation with 10 realizations.

	Supine5	Early5	Last5	Early15	Early15
Parameters	LFnu, HF, LF/HF, $mean_RR$	LF, LFnu, LF/HF, $mean_RR$	LF, P_{tot} , $mean_RR$, sd_RR	P_{tot} , sd_RR , $\alpha_2(RR)$, $SampEn(RR)$	HF, HFnu, $\alpha_2(RR)$, BRS_{Up}
σ	0.36	0.48	0.032	0.57	0.22
Sensitivity	65%	61.1%	65.1%	80.8%	98.8%
Specificity	62%	58.7%	50.3%	62.3%	30.2%

- **Feature selection methods**

All the parameters computed during Early15 were evaluated using the two different feature selection methods described in chapter 3 (relief method and *probe* feature), in

order to extract the most relevant descriptors that help in the classification, table 4.4 summarizes the result of relief method with 1 is the first and 19 is the last parameter ranked by relief, and the figure 4.9 represents the probability distribution functions that the *probe* feature is better classified than a variable candidate of the set of parameters computed in Early15.

TAB. 4.4- Results of relief method with 1 is the first and 19 is the last ranking parameter

ranking by relief method	Parameters
1	SampEn(RR)
2	sd_RR
3	α_2 (RR)
4	SampEn(Δ RR)
5	P_tot
6	mean_RR
7	pRR50
8	LFnu
9	HFnu
10	LF/HF
11	α_1 (RR)
12	BRS
13	HF
14	LF
15	rmssd
16	α_2 (Δ RR)
17	DownBRS
18	UpBRS
19	α_1 (Δ RR)

We noticed that the two feature selection methods gave mainly the same ranking for the four parameters that provided the best result in the exhaustive search. Table 4.5 shows the average result of each 10 iterations of classification using combination between one of the feature selection algorithms (relief method or *probe* feature) and one of the classification methods (KSVM or KNN).

- **Performance with correct classification of patients with positive response.** From clinical point of view, it is fundamental to correctly classify all the positive subjects. So, an exhaustive procedure was also applied in order to correctly classify all patients with positive response using LOOCV; the results show that in Supine5, Early5 and Last5 a 100% correct classification of positive test group was not reached. There were respectively 6%, 18.2% and 18.2% of patients with positive response incorrectly classified with global error (error of correctly classify negative and positive test groups)

of 59.3% in Supine5, 56.1% in Early5 and 56.1% in Last5. In Early15, while using the four parameters HF, HFnu, α_2 (RR) and BRS_{Up} with $\sigma=0.22$ as radial basis kernel parameter, we reach 100% correct classification of positive patients with an accuracy of the classifier equal to 72.1%, and a limited specificity of 39.2%.

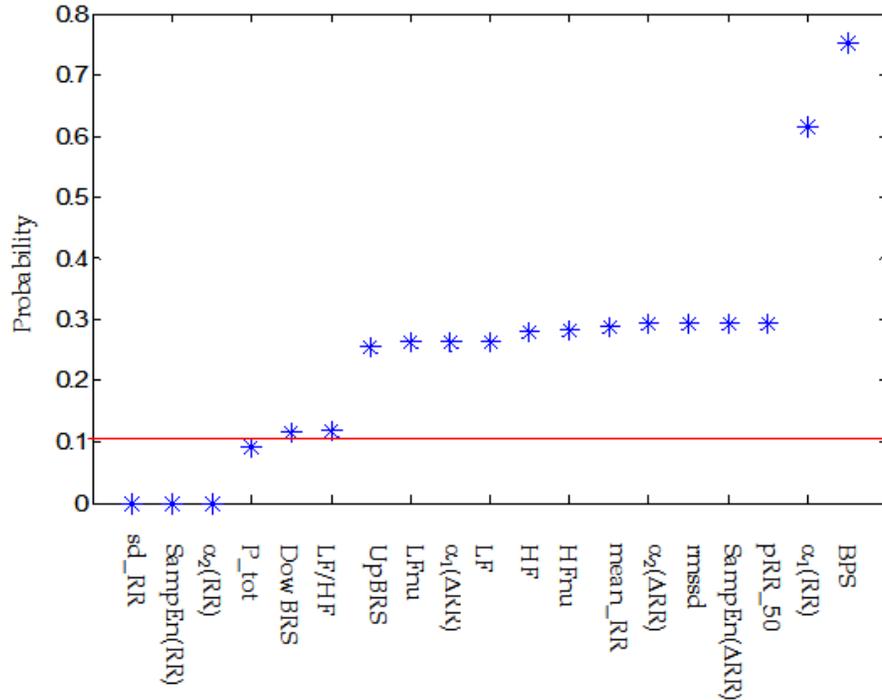


FIG. 4.9 - The probability distribution functions that the 'probe' feature is better classified than a variable candidate of the set of parameters computed in Early15. Parameters with probability<0.1 (under the red line) are the most pertinents.

TAB. 4.5- The average results of 10 iterations of classification using combinations presented in the table.

	Sensitivity	Specificity
Relief/KNN	65.4%	63.5%
Relief/KSVM	70.3%	74.2%
Sonde/KNN	64.3%	65.2%
Sonde/KSVM	80.8%	62.3%

4.3. Discussion

In intra-group comparison, reciprocal changes occurred in the sinus node response to sympathetic and parasympathetic activity when passing from supine to tilted position in subjects with positive and negative responses, which was reflected by the increase in the LFnu and the decrease in the HFnu. Such results suggest a reduction of the sinus node response to parasympathetic activity with an increased response to sympathetic activity in response to orthostatic stress due to HUTT (Moya *et al* 2009).

These results are in accordance with the well-known early autonomic response to HUTT already reported in the literature (Pagani *et al* 1986, Montano *et al* 1994, Vybiral *et al* 1989).

In inter-group comparison, using 5-minute long windows, only the LF parameter showed a significant difference between the two groups in Early5 and Last5. None of the other classical parameters provided a significant difference between groups during the supine position. Nevertheless, using 15 minutes of the tilting position, we clearly observed that the syncope positive HUTT group had higher sd_RR than the negative group. This 'hyperactivation' reflected by the high value of sd_RR may be explained by the fact that in the positive group, the reflex seems insufficiently efficient (LF/HF ratio goes from 1.57 as mean value at rest to 4.9 early after orthostatism, whereas in the negative group LF/HF ratio goes from 1.33 to 8.24), so the adaptation response appears permanently activated in order to try to balance HR and pressure. This can be linked to the Bezold-Jarisch reflex which is responsible for vaso-vagal syncopes and occurs when (1) preload is low and (2) sympathetic tone, conducted by hormones (catecholamine) and nervous messengers to the sinus node, is high (Mosqueda-Garcia *et al* 2000). The parameters α_2 (RR) and α_2 (Δ RR) are higher and SampEn(RR) is lower in the positive than in the negative group. This shows an increase in long term correlation properties (Peng *et al* 1995) and a decrease in overall complexity of heart rate dynamics in the positive group. Therefore, by enlarging the window length in the tilting position, an early significant difference between groups can be found.

In comparison with previous studies, our results are in agreement with the findings of (Kim *et al* 2000), who noted that entropies decreased in positive test group with neurocardiogenic syncope compared with patients with negative HUTT. Moreover, in healthy subjects, (Tulppo *et al* 2001) noted that α_1 (RR) increased and ApEn and α_2 (RR) did not change during HUTT. However, they did not mention if the study population had negative or positive tilt test. By contrast, (Kubo *et al* 2003) reported a decrease in ApEn and an increase in α_2 (RR) in their subject population which had negative HUTT. Both studies evaluated the effect of HUTT on the dynamics of heart rate without comparing the difference among subjects with negative or positive response.

Furthermore, the proposed approach based on KSVM showed a discriminative classification that could classify correctly 85 % of patients after 15 min of HUTT, using linear and non-linear parameters. This implies that only 15 % of subjects can be misidentified from the first 15 minutes from the beginning of the tilting test. This shows that the heart rate dynamics during the 15 minutes of HUTT are the more relevant in

terms of discrimination between the two groups. Thus, in order to predict HUTT result (i.e. positive or negative), the first 15 min seem more relevant than the first 5 minutes previously analyzed in some studies (Kouakam *et al* 1999). Moreover, the first 15 min seem also more relevant in terms of classification than the 5 min just before the syncope. Such results are in agreement with the fact that syncope mainly occurs between 15 to 25 min of HUTT (Fitzpatrick *et al* 1991). Our results tend to confirm that the autonomic adaptations are not efficient during the first 15 min of a HUTT in subjects experiencing syncope. These findings could be explained by two main aspects: *i*) a higher sympathetic tone with a lack of adaptation since the beginning of the HUTT, leading to syncope in positive subjects and *ii*) a blunt of sympathetic tone in the last 5 minutes of HUTT, due to the drop in HR and peripheral resistances, linked to huge parasympathetic tone. The obtained results might be improved by enlarging the study population, taking the same number of subjects for both the training and testing phases. The possibility of identifying patients with positive response to HUTT before syncope occurs is of interest from a physiological viewpoint, motivating the reduction of the clinical time of HUTT. In addition, reflex syncope mechanisms are still poorly understood. Analyzing HRV in supine posture and tilt posture in asymptomatic subjects could help to better understand syncope mechanisms. This could improve syncope diagnosis and care.

Previous studies have suggested an early prediction of the tilt test outcomes (Mallat *et al* 1997, Bellard *et al* 2001, Schang *et al* 2007, Gimeno-Blanes *et al* 2011). The original aspects of our study are (1) the population included in this work, containing only asymptomatic subjects, (2) the use of non-linear parameters that have shown interesting results in previous studies and (3) the applied signal processing and classification methods, with varying time support that enabled the identification of the information content of the Early15 phase. Indeed, Comparing with other similar works in the literature, this larger analysis window (15 min) allows for the computing of the parameters with more data. Thus they can provide more interesting information without the need to extend the testing duration and can lead to the reduction of the examination time. Our study provides interesting results with a performance (Se=88.5% and Sp=80.6%) and that can be considered higher than others reported in the literature (Sumiyoshi *et al.*, 2000), (Bellard *et al.*, 2001), (Movahed and Hassapoyannes, 2001), (Pitzalis *et al.*, 2002). In addition, the proposed approach provides these performance figures while analyzing data from the first minutes after the beginning of HUTT. Indeed, one of the most extensive studies in the literature (the largest database) (Virag *et al.*, 2007) yielded 95% sensitivity and 93 % specificity using a combination of HR and BP features, but 51% of the results were based on data from the last minute before syncope.

4.4. Conclusion

Our results show that using 15-minute wide windows in the tilting position can lead to early distinction between subjects with negative response and others with positive response to HUTT. We have found that non-linear parameters extracted from RR and SS series during the Early15 phase significantly differ between both groups. Compared to positive subjects, those with negative HUTT show lower standard deviation and long-term fractal scale of RR-interval time series, and greater sample entropy during the tilted position. This allowed us to construct a model using KSVM that classifies patients into negative and positive groups, by analyzing the first 15 minutes of the tilted position with 85% of performance (Sensitivity=88.5% and Specificity=80.6%) using together the four following parameters: P_{tot} , sd_{RR} , $SampEn(RR)$ and $\alpha_2(RR)$. The application of the proposed method with its possibility to have a correct classification of all positive subjects during classical HUTT can have a large interest in the clinical practice, helping also to reduce the examination time.

Finally, the use of a new database containing patients suffering for unexplained syncope with positive and negative responses to HUTT to validate our results (validation set) is possible in the future. Moreover, it would be interesting to complete this classification using relationships between ECG parameters (such as RR-interval time series), blood pressure information (such as systolic-to-systolic time series) and respiration indexes in time-frequency domain (Orini *et al* 2012). This is the topic of the next chapter.

Appendix A

Detrended Fluctuation Analysis and Sample Entropy

Detrended Fluctuation Analysis. First, the centered series of RR interval $s(k)$ of length L undergoes cumulative summing to obtain the integrated series $y(k)$:

$$y(k) = \sum_{i=1}^k [s(i) - s_{av}] \quad (A1)$$

With

$$s_{av} = \frac{1}{L} \sum_{k=1}^L s(k) \quad (A2)$$

The integrated series is then decomposed into segments of equal length n . In each segment, the local trend $y_n(k)$, obtained by a least-squares line fit, is subtracted from the integrated series to detrend the data. The fluctuation $F(n)$ for a segment of length n was calculated as the root mean square of the integrated and detrended segment.

$$F(n) = \sqrt{\frac{1}{L} \sum_{k=1}^L [y(k) - y_n(k)]^2} \quad (A3)$$

This calculation is repeated for different segment sizes (n). A linear relationship can be typically found between n and $F(n)$ in a logarithmic scale. The scaling exponent α defines the slope of the line which relates $\log F(n)$ to $\log n$ thus it express the autocorrelation characteristics of the signal:

- $\alpha < 0.5$: anti-correlated signal
- $\alpha = 0.5$: uncorrelated signal
- $\alpha > 0.5$: autocorrelation in the signal

Based on literature for heart rate measurements, the slope is generally calculated in two ranges of the segment to obtain the short and long range scaling exponents α_1 and α_2 respectively.

Sample Entropy. For a data series $x(n)=x(1),x(2),\dots,x(L)$, where L is the length of the series, two parameters are defined: the embedded dimension m of the vector to be formed and r , a threshold that is used as noise filter. The procedure to compute Sample entropy (SampEn) is shown in the following steps:

1) $L-m+1$ vectors of size m each vector, are composed as:

$$X^m(i) = [x(i), x(i+1), \dots, x(i+m-1)] \quad \text{With } i = 1, L - m + 1$$

- 2) The distance between each vector and the other vector (including itself) is computed as:

$$d[X^m(i), X^m(j)] = \max \left[|x(i+k) - x(j+k)| \right] \text{ With } 0 \leq k \leq m-1 \quad (\text{A4})$$

- 3) For each vector $X^m(i)$, the number of template matching:

$$N^m(i) = \text{card} \{ d[X^m(i), X^m(j)] \leq r \}$$

And $C_r^m(i) = N^m(i)/(L-m+1)$ is the probability that any vector $X^m(j)$ matches

$X^m(i)$. The average of $C_r^m(i)$ over i is:

$$\phi^m(r) = \frac{1}{L-m+1} \sum_{i=1}^{L-m+1} C_r^m(i) \quad (\text{A5})$$

- 4) By increasing the size of the vector to $m+1$ and repeating the previous step to obtain $\phi^{m+1}(r)$, the theoretical value of SampEn is:

$$\text{SampEn}(m, r) = \lim_{L \rightarrow \infty} [\ln(\phi^m(r) - \phi^{m+1}(r))] \quad (\text{A6})$$

For a finite value of L , the estimate of the SampEn is defined as:

$$\text{SampEn}(m, r, L) = \ln[(\phi^m(r) - \phi^{m+1}(r))] \quad (\text{A7})$$

References

- Barquero-Perez O, de Sa J M, Rojo-Alvarez J L and Goya-Esteban R 2008 Changes in Detrended Fluctuation indices with aging in healthy and Congestive Heart Failure subjects, in: Computers in Cardiology 2008. Presented at the Computers in Cardiology 2008 pp. 45–48. doi:10.1109/CIC.2008.4748973
- Bellard E, Fortrat J-O, Vielle B, Dupuis J-M, Victor J and Lefthériotis G 2001 Early predictive indexes of head-up tilt table testing outcomes utilizing heart rate and arterial pressure changes Am. J. Cardiol. 88 903–906. doi:10.1016/S0002-9149(01)01904-X
- Brignole M, 2007 Diagnosis and treatment of syncope Heart 93 130–136. doi:10.1136/hrt.2005.080713
- Brignole M, Menozzi C, Del Rosso A, Costa S, Gaggioli G, Bottoni N, Bartoli P and Sutton R 2000 New classification of haemodynamics of vasovagal syncope: beyond the VASIS classification. Analysis of the pre-syncopal phase of the tilt test without and with nitroglycerin challenge. Vasovagal Syncope International Study. Eur. Pacing Arrhythm. Card. Electrophysiol. J. Work. Groups Card. Pacing Arrhythm. Card. Cell. Electrophysiol. Eur. Soc. Cardiol. 2 66–76.

- Dumont J, Hernandez A I and Carrault G 2010 Improving ECG Beats Delineation With an Evolutionary Optimization Process *IEEE Trans. Biomed. Eng.* 57 607–615. doi:10.1109/TBME.2008.2002157
- Electrophysiology T F of the E S of C the N A S of P 1996 Heart Rate Variability Standards of Measurement Physiological Interpretation and Clinical Use *Circulation* 93 1043–1065. doi:10.1161/01.CIR.93.5.1043
- Fitzpatrick A P, Theodorakis G, Vardas P and Sutton R 1991 Methodology of head-up tilt testing in patients with unexplained syncope *J. Am. Coll. Cardiol.* 17 125–130.
- Francis D P, Willson K, Georgiadou P, Wensel R, Davies L C, Coats A and Piepoli M 2002 Physiological basis of fractal complexity properties of heart rate variability in man *J. Physiol.* 542 619–629.
- Gimeno-Blanes F J, Rojo-Álvarez J L, Caamaño A J, Flores-Yepes J A and García-Alberola A 2011. On the feasibility of tilt test outcome early prediction using ECG and pressure parameters *EURASIP J. Adv. Signal Process.* 2011 33 doi:10.1186/1687-6180-2011-33
- Goldberger A L, Amaral L A N, Glass L, Hausdorff J M, Ivanov P C, Mark R G, Mietus J E, Moody G B, Peng C-K and Stanley H E 2000 PhysioBank PhysioToolkit and PhysioNet Components of a New Research Resource for Complex Physiologic Signals *Circulation* 101 e215–e220. doi:10.1161/01.CIR.101.23.e215
- Goya-Esteban R, de Sa J P M, Rojo-Alvarez J L and Barquero-Perez O 2008 Characterization of Heart Rate Variability loss with aging and heart failure using Sample Entropy, in: *Computers in Cardiology 2008 Presented at the Computers in Cardiology 2008* pp. 41–44. doi:10.1109/CIC.2008.4748972
- Holmegard H N, Benn M, Kaijer M, Haunsø S and Mehlsen J 2012 Differences in autonomic balance in patients with cardioinhibitory and vasodepressor type of reflex syncope during head-up tilt test and active standing *Scand. J. Clin. Lab. Invest.* 72 265–273. doi:10.3109/00365513.2012.659282
- Karapetian G K, Engels H J, Gretebeck K A and Gretebeck R J 2012 Effect of caffeine on LT, VT and HRVT *Int. J. Sports Med.* 33 507–513. doi:10.1055/s-0032-1301904
- Kim J S, Park J E, Seo J D, Lee W R, Kim H S, Noh J I, Kim N S and Yum M K 2000 Decreased entropy of symbolic heart rate dynamics during daily activity as a predictor of positive head-up tilt test in patients with alleged neurocardiogenic syncope *Phys. Med. Biol.* 45 3403–3412.
- Kouakam C, Lacroix D, Zghal N, Logier R, Klug D, Le Franc P, Jarwe M and Kacet S 1999 Inadequate sympathovagal balance in response to orthostatism in patients with unexplained syncope and a positive head up tilt test *Heart* 82 312–318.

- Kubo Y et al. 2003 Toward chronocardiologic and chronomic insights: dynamics of heart rate associated with head-up tilting. *Biomed. Pharmacother. Bioméd. Pharmacothérapie* 57 Suppl 1, 110s–115s.
- Lake D E, Richman J S, Griffin M P and Moorman J R 2002 Sample entropy analysis of neonatal heart rate variability *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 283 R789–797. doi:10.1152/ajpregu.00069.2002
- Lippman N, Stein K M and Lerman B B 1995 Failure to decrease parasympathetic tone during upright tilt predicts a positive tilt-table test *Am. J. Cardiol.* 75 591–595.
- Madrid A H, Moro C, Marín-Huerta E, Novo L, Mestre J L, Lage J and Ricoy E 1994 [Usefulness of the RR variability in the diagnosis of neurogenic syncope] *Rev. Esp. Cardiol.* 47 536–543.
- Mallat Z, Vicaut E, Sangaré A, Verschueren J, Fontaine G and Frank R 1997 Prediction of head-up tilt test result by analysis of early heart rate variations *Circulation* 96 581–584.
- Montano N, Ruscone T G, Porta A, Lombardi F, Pagani M and Malliani A 1994 Power spectrum analysis of heart rate variability to assess the changes in sympathovagal balance during graded orthostatic tilt *Circulation* 90 1826–1831.
- Mosqueda-Garcia R, Furlan R, Tank J and Fernandez-Violante R 2000 The elusive pathophysiology of neurally mediated syncope *Circulation* 102 2898–2906.
- Movahed M R and Hassapoyannes C A 2001 Prediction of non-occurrence of syncope during a tilt-table test by early heart rate variations *J. S. C. Med. Assoc.* 1975 97 207–210.
- Moya A et al. 2009 Guidelines for the diagnosis and management of syncope (version 2009) The Task Force for the Diagnosis and Management of Syncope of the European Society of Cardiology (ESC) *Eur. Heart J.* 30 2631–2671. doi:10.1093/eurheartj/ehp298
- Orini M, Laguna P, Mainardi L T and Bailón R 2012 Assessment of the dynamic interactions between heart rate and arterial pressure by the cross time-frequency analysis *Physiol. Meas.* 33 315–331. doi:10.1088/0967-3334/33/3/315
- Pagani M, Lombardi F, Guzzetti S, Rimoldi O, Furlan R, Pizzinelli P, Sandrone G, Malfatto G, Dell’Orto S and Piccaluga E 1986 Power spectral analysis of heart rate and arterial pressure variabilities as a marker of sympatho-vagal interaction in man and conscious dog. *Circ. Res.* 59 178–193.
- Parry S W and Tan M P 2010 An approach to the evaluation and management of syncope in adults *BMJ* 340 c880–c880. doi:10.1136/bmj.c880
- Peng C K, Havlin S, Stanley H E and Goldberger A L 1995 Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series *Chaos Woodbury N* 5 82–87. doi:10.1063/1.166141

- Pincus S M 1991 Approximate entropy as a measure of system complexity Proc. Natl. Acad. Sci. U. S. A. 88 2297–2301.
- Pitzalis M, Massari F, Guida P, Iacoviello M, Mastropasqua F, Rizzon B, Forleo C and Rizzon P 2002 Shortened head-up tilting test guided by systolic pressure reductions in neurocardiogenic syncope Circulation 105 146–148.
- Richman J S and Moorman J R 2000 Physiological time-series analysis using approximate entropy and sample entropy Am. J. Physiol. Heart Circ. Physiol. 278 H2039–2049.
- Schang D, Feuilloy, M, Plantier G, Fortrat J-O and Nicolas P 2007 Early prediction of unexplained syncope by support vector machines Physiol. Meas. 28 185. doi:10.1088/0967-3334/28/2/007
- Shawe-Taylor J and Cristianini N 2004 Kernel Methods for Pattern Analysis Cambridge University Press.
- Sumiyoshi M, Nakata Y, Mineda Y, Tokano T, Yasuda M, Nakazato Y and Yamaguchi H 2000 Does an early increase in heart rate during tilting predict the results of passive tilt testing? Pacing Clin. Electrophysiol. PACE 23 2046–2051.
- Sutton R, Petersen M, Brignole M, Raviele A and Menozzi C 1992 Proposed classification for tilt induced vasovagal syncBellard, E., Fortrat, J.O., Vielle, B., Dupuis, J.M., Victor, J., Lefthériotis, G., 2001. Early predictive indexes of head-up tilt table testing outcomes utilizing heart rate and arterial pressure changes. Am. J. Cardiol. 88, 903–906, A8.
- Bertinieri, G., di Rienzo, M., Cavallazzi, A., Ferrari, A.U., Pedotti, A., Mancia, G., 1985. A new approach to analysis of the arterial baroreflex. J. Hypertens. Suppl. Off. J. Int. Soc. Hypertens. 3, S79–81.
- Deharo, J.-C., Jegou, C., Lanteaume, A., Djiane, P., 2006. An implantable loop recorder study of highly symptomatic vasovagal patients: the heart rhythm observed during a spontaneous syncope is identical to the recurrent syncope but not correlated with the head-up tilt test or adenosine triphosphate test. J. Am. Coll. Cardiol. 47, 587–593. doi:10.1016/j.jacc.2005.09.043
- Gouveia, S., Rocha, A., Van de Borne, P., Lago, P., 2005. Assessing baroreflex sensitivity in the sequences technique: local versus global approach, in: Computers in Cardiology, 2005. Presented at the Computers in Cardiology, 2005, pp. 279–282. doi:10.1109/CIC.2005.1588091
- Movahed, M.R., Hassapoyannes, C.A., 2001. Prediction of non-occurrence of syncope during a tilt-table test by early heart rate variations. J. S. C. Med. Assoc. 1975 97, 207–210.

- Pitzalis, M., Massari, F., Guida, P., Iacoviello, M., Mastropasqua, F., Rizzon, B., Forleo, C., Rizzon, P., 2002. Shortened head-up tilting test guided by systolic pressure reductions in neurocardiogenic syncope. *Circulation* 105, 146–148.
- Pitzalis, M.V., Mastropasqua, F., Passantino, A., Massari, F., Ligurgo, L., Forleo, C., Balducci, C., Lombardi, F., Rizzon, P., 1998. Comparison Between Noninvasive Indices of Baroreceptor Sensitivity and the Phenylephrine Method in Post-Myocardial Infarction Patients. *Circulation* 97, 1362–1367. doi:10.1161/01.CIR.97.14.1362
- Sutton, R., Bloomfield, D.M., 1999. Indications, methodology, and classification of results of tilt-table testing. *Am. J. Cardiol.* 84, 10–19. doi:10.1016/S0002-9149(99)00692-X
- Virag, N., Sutton, R., Vetter, R., Markowitz, T., Erickson, M., 2007. Prediction of vasovagal syncope from heart rate and blood pressure trend and variability: experience in 1,155 patients. *Heart Rhythm Off. J. Heart Rhythm Soc.* 4, 1375–1382. doi:10.1016/j.hrthm.2007.07.018
- Tulppo M P, Hughson R L, Mäkikallio T H, Airaksinen K E, Seppänen T and Huikuri H V 2001 Effects of exercise and passive head-up tilt on fractal and complexity properties of heart rate dynamics *Am. J. Physiol. Heart Circ. Physiol.* 280 H1081–1087.
- Vybiral T, Bryg R J, Maddens M E and Boden W E 1989 Effect of passive tilt on sympathetic and parasympathetic components of heart rate variability in normal subjects *Am. J. Cardiol.* 63 1117–1120.
- Khodor, N., Matelot, D., Carrault, G., Amoud, H., Khalil, M., Ville, N., Carre, F., Hernandez, A., 2014. Kernel based support vector machine for the early detection of syncope during head-up tilt test. *Physiol. Meas.* 35, 2119–2134. doi:10.1088/0967-3334/35/10/2119
- Paola Martínez-García, C.L., 2012. Relation of the baroreflex mechanism with the photoplethysmographic volume in healthy humans during orthostatism. *Arch. Cardiol. Mex.* 82, 82–90.

Chapter 5

Early syncope detection during head-up tilt test by analyzing interactions between cardio-vascular signals

The evaluation of the dynamic interactions between several cardiovascular variables can play an important role to improve our understanding on the function of the baroreflex control of arterial blood pressure and can help to enhance the prediction performance. For this aim, the smoothed pseudo Wigner-Ville distribution (*SPWD*) approach was used in this chapter, to characterize the time-frequency relationship between time series extracted from ECG (*RR*-interval) and blood pressure (amplitudes of the systolic blood pressure, the peak amplitude of the first derivative of the blood pressure, dP/dt and the variability of the pulse transit time, *PTT*). Two new indexes quantifying the relationship between time series were introduced. The first index computed the distribution of time-frequency coherence between 2 time series in different frequency bands and the second one computed the scalar product of the time-frequency spectra of these time series.

In this chapter, an overview of the recent studies that have used time-frequency methods in biomedical applications are first reported, then the main characteristics of the analysis and the preprocessing steps are recalled. After that in time-frequency analysis method paragraph, the analysis method based on the *SPWD* coherence are described, then the two features set extracted from the coherence are presented. As we proposed a set of selection descriptors, several feature selection methods and classification methods are compared. The results are reported in the results paragraph. In addition, a comparison study was carried out between the *SPWD* method and a method that computes the linear correlation coefficient in the time and frequency domains. The last section is focused on a large discussion that concludes this contribution.

5.1. General overview

Significant temporal, structural and spectral changes occur during the HUTT procedure, leading to highly non-stationary observations. Such data justifies the interest of using non-stationary techniques such as time-frequency (TF) analysis, rather than stationary analysis methods. We hypothesize that using such methods would provide additional information that may be useful for improving the diagnosis of vasovagal syncope and to improve our understanding of the cardiovascular dysfunctions at work in fainting subjects.

TF methods have been widely applied to biomedical signal analysis, particularly in the study of cardiac control. A review of the prevalent tools for TF analysis can be found in (Mainardi et al., 2002). Relationships in the time-frequency plane for non-stationary signals have been also investigated in different areas. For instance, a new index was proposed in (Ansari-Asl et al., 2005) to characterize the interdependency between different EEG channels, by estimating a linear correlation coefficient in the time and frequency domains. More recently, this method has been applied to the diagnosis of infection in premature infants (Carrault et al., 2009). Moreover, in 2012, (M. Orini et al., 2012) have evaluated the dynamic interaction between HR and BP using the cross-TF analysis, based on the smoothed pseudo-Wigner-Ville distribution (*SPWD*). They have demonstrated that HUTT induces an abrupt reduction in the cardiovascular coupling in the high frequency band for healthy patients.

Based on the interesting results already reported (M Orini et al., 2012), the *SPWD* has been employed in this chapter, to evaluate the dynamic interaction between RR-interval and several parameters issued from systolic blood pressure signal to distinguish between patients that would develop a syncope or not, before syncope or pre-syncope symptoms occur. Briefly, our approach can be described by the block diagram depicted on figure 5.1:

1. After ECG and Pressure signals acquisition during the HUTT, the signals undergo a processing step to extract different time series. Heart rate variability signal (*RR-interval*) is classically extracted from the ECG signal while the amplitude of the systolic blood pressure (*AmpS*), the peak magnitude of the first derivative of systolic blood pressure (*dPdt_max*), and the variability of the pulse transit time (*PTT*) are derived from the systolic blood pressure.
2. The *SPWD* coherence between these time series is then computed and parameters are extracted from the *SPWD* coherence.

3. These parameters go through a feature selection (FS) algorithm to choose the relevant indicators to the classification process (dotted line indicates the joint task between feature selection and the classifier).
4. Finally, the output of this model is obtained as a patient's classification with positive (will develop syncope during HUTT) or negative response (will not develop syncope during the test).

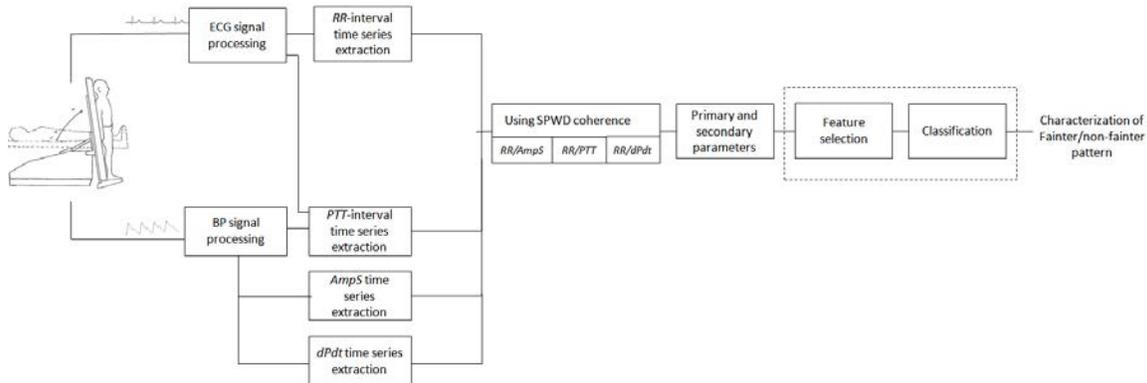


FIG. 5.1- Block diagram of the proposed approach for syncope prediction.

5.2. Data acquisition and signal processing

Here 9 subjects are excluded because their improper blood pressure signals to have 57 subjects composed from 29 who did not present syncope and thus were considered as non-fainters (negative subjects), and 28 who developed vasovagal syncope and were assigned as fainters (positive subjects).

The signal processing was reported in chapter 3 (section 3.2) where different time series were extracted and used in this chapter: the amplitude of the systolic wave (*AmpS*), the Pulse Transit Time (*PTT*) and the peak amplitude of the first derivative of the blood pressure signal (*dPdt_{max}*).

These time-series were extracted on the following time intervals of HUTT, as depicted on figure 5.2 and figure 5.3:

- During 5 minutes of the supine period with controlled breathing (Supine5).
- During the first 5 minutes of the HUTT (Early5).
- During the first 15 minutes of the HUTT (Early15)
- During the last 5 minutes of the HUTT (Last5):
 - Before syncope occurs; if the patient was fainter.
 - Before ending the HUTT; if the patient was non-fainter.

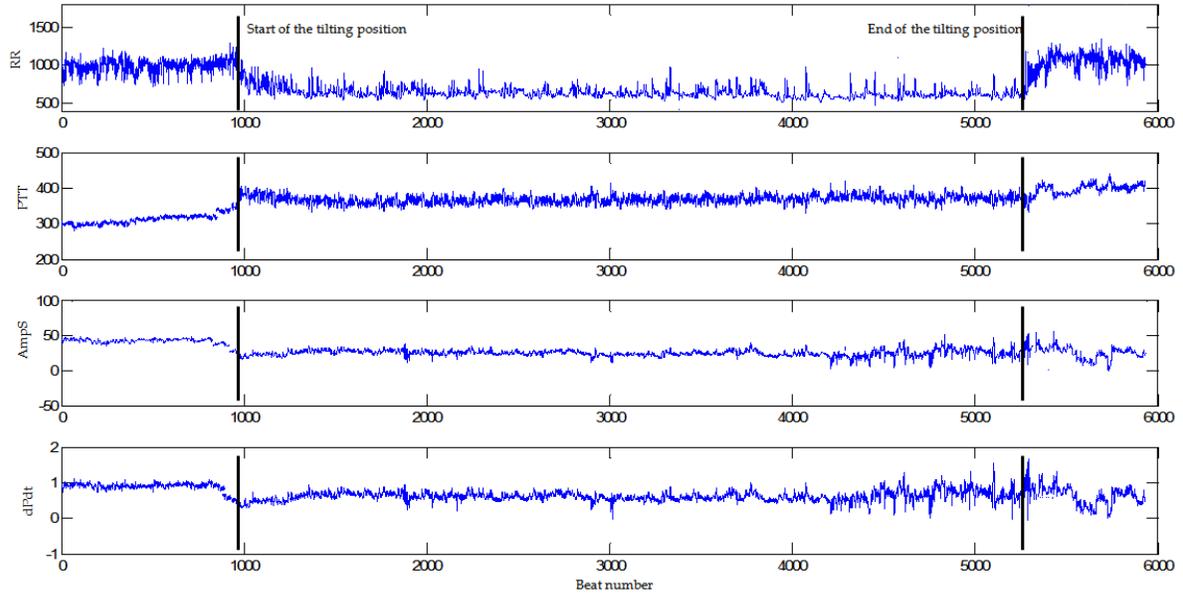


FIG. 5.2 - Example of the full time series extracted from ECG and blood pressure of subject with negative response to HUTT

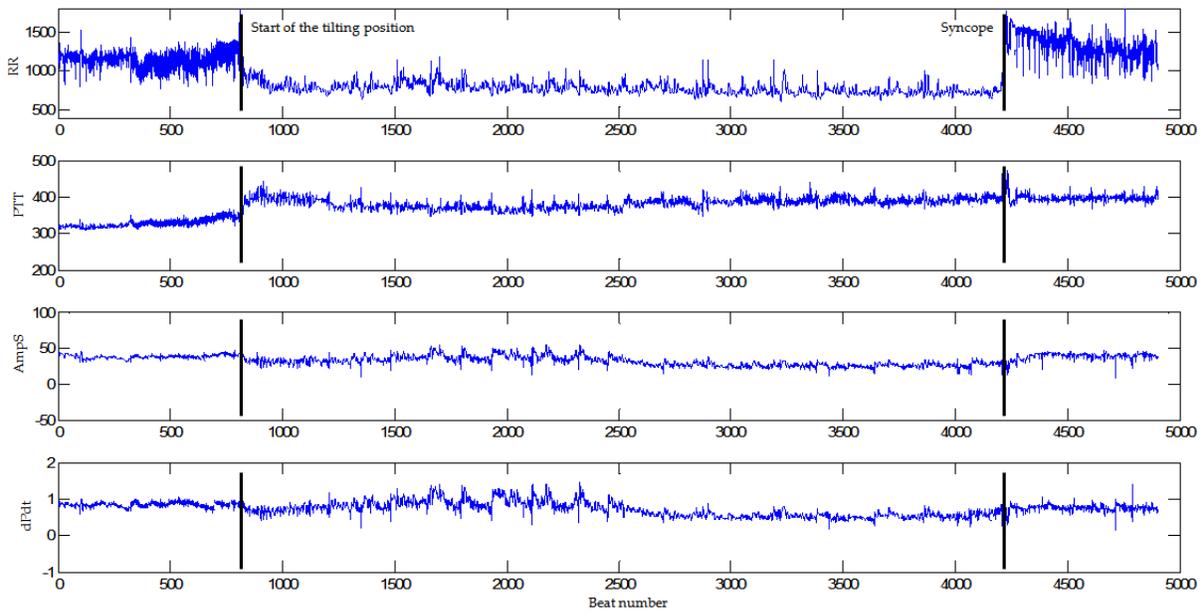


FIG. 5.3 - Example of the full time series extracted from ECG and blood pressure of subject experiencing syncope

Figure 5.4 and 5.5 represents the time series extracted from a subject with negative response to HUTT and other who experienced syncope during the different conditions cited above. These figures illustrate the non stationary of the data and justify the interest to use time-frequency methods.

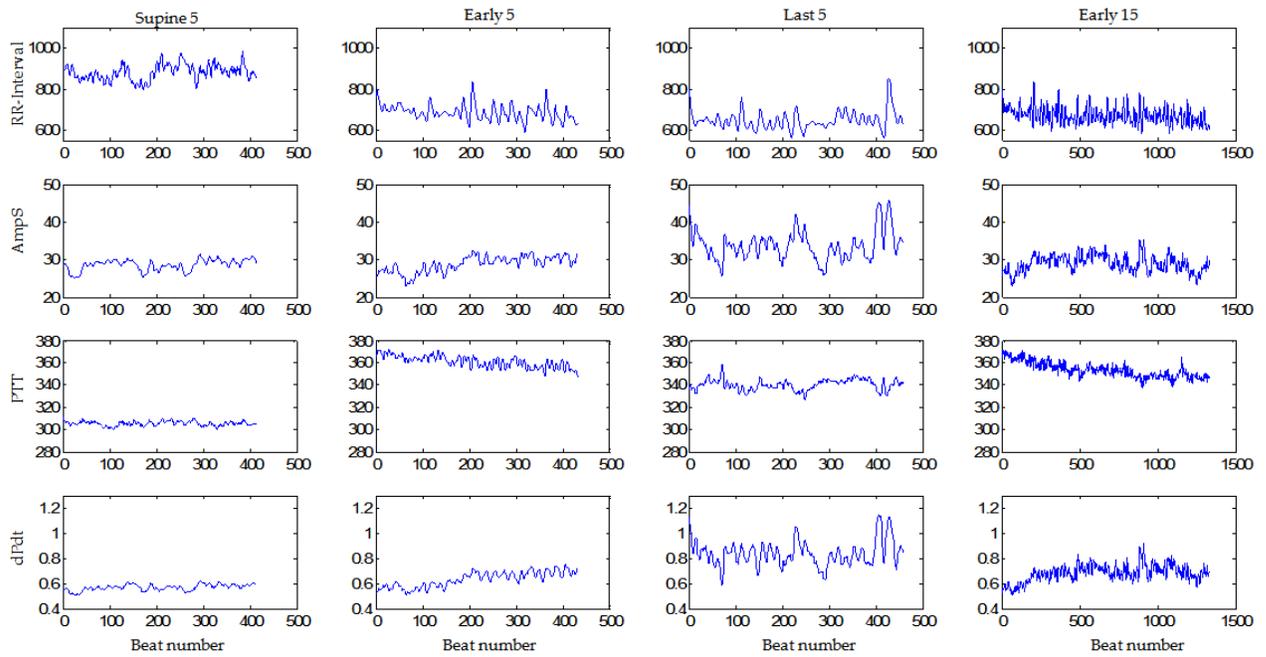


FIG. 5.4 - Different time series during the conditions cited above, extracted from a subject with negative response to HUTT.

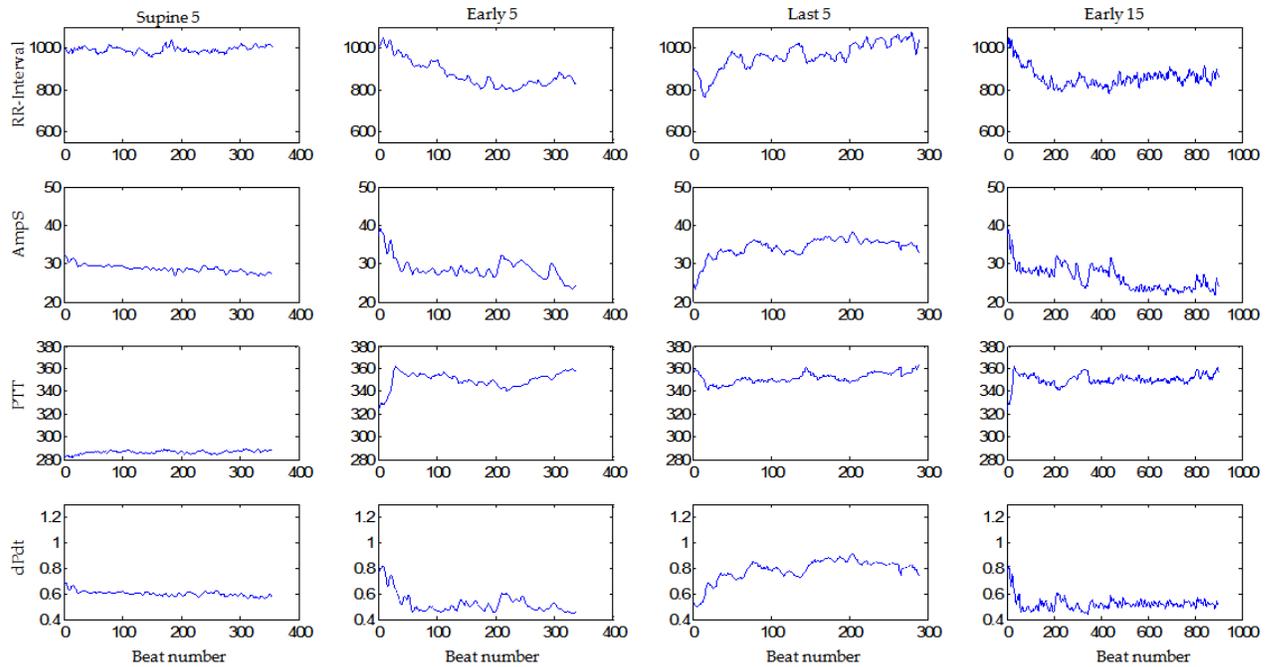


FIG. 5.5 - Different time series during the conditions cited above, extracted from a subject who experienced syncope.

5.3. Time-Frequency analysis method

5.3.1. Time Frequency coherence using SPWD

The Time Frequency coherence computes the local coupling between two signals in TF domain, and can be defined as (Matz and Hlawatsch, 2000) (White and Boashash, 1990):

$$\gamma(t, f) = \frac{|C_{xy}(t, f)|}{\sqrt{C_{xx}(t, f)C_{yy}(t, f)}} \quad (5.1)$$

Where $C_{xy}(t, f)$ is the cross TF spectrum between $x(t)$ and $y(t)$, $C_{xx}(t, f)$ and $C_{yy}(t, f)$ are the auto TF spectrum of $x(t)$ and $y(t)$ respectively. Here, the SPWD is used to estimate the auto and cross TF spectra, as follows:

$$C_{xy}(t, f) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \phi(\tau, v) A_{xy}(\tau, v) e^{j2\pi(tv - tf)} dv d\tau \quad (5.2)$$

$$A_{xy}(\tau, v) = \int_{-\infty}^{+\infty} x\left(t + \frac{\tau}{2}\right) y^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi vt} dt \quad (5.3)$$

Where $A_{xy}(\tau, v)$ is the cross-ambiguity function of the two signals $x(t)$ and $y(t)$, $\Phi(\tau, v)$ is the kernel. In order to have physical interpretation $\gamma(t, f)$ should be between 0 and 1. Therefore, one should suppress the interference (Matz and Hlawatsch, 2000). To do this, an elliptical exponential function is used as a kernel.

$$\phi(\tau, v) = \exp \left\{ -\pi \left[\left(\frac{v}{v_0} \right)^2 + \left(\frac{\tau}{\tau_0} \right)^2 \right]^{\lambda} \right\} \quad (5.4)$$

The degree of frequency and time filtering are defined by τ_0 and v_0 respectively. They set the length of the ellipse axes of $\phi(\tau, v)$ isocontours along τ and v . The parameter λ defines the size of the extremities of the kernel. The convenient degree of smoothing is obtained by setting an appropriate TF resolution. Temporal and spectral resolutions are estimated using distinctive combinations of time and frequency resolutions respecting the condition $\gamma(t, f) \in [0, 1]$; the most appropriate combination is that with smaller values.

5.3.2. Feature extraction

Several indexes are computed from the SPWD coherence for the classification of patients with positive or negative response and the early prediction of the HUTT result. Briefly, they can be divided in two categories: the distribution of time-frequency coherence over a threshold, and the scalar product of the TF spectra.

5.3.2.1. Distribution of time-frequency coherence over a threshold

To compute the distribution of time-frequency coherence is equivalent to enumerating the number of values of the time frequency coherence (TFC) that are higher than threshold $[\gamma^{band}(t,f)]_{th}$ in a specified bandwidth and over a temporal window; this gives an idea of the TFC repartition.

$$[\gamma^{band}(t,f)]_{th} = [\text{number of TFC values} > th]_{band}$$

Where th is the threshold value and $band$ is the sub-band under study. The selection of the bands and the tuning of the threshold for this particular application are detailed in the results section. They are classically selected from physiopathological knowledge such as the sympathetic and parasympathetic influences on the spectra.

5.3.2.2. The scalar product of the TF spectra

The scalar product of the TF spectra index has a definite physical explanation. It reveals the similarity between signals. Higher scalar products are related to a higher similarity between signals in TF domain. By defining W_R and W_P the normalized *SPWD* of RR-interval time series and one of the time series extracted from blood pressure or the *PTT* time series, the scalar product S_{band} of the auto TF spectrum of the two signals, computed in a specific band, is defined as

$$S_{band} = \langle W_R, W_P \rangle_{band}$$

5.3.2.3. Secondary indexes

It is well-known that new parameters can be derived based on the primary indexes described above (Dreyfus et al., 2011). Indeed, simple combinations (product and/or division) between primary indexes can be statistically more significant and help to reduce the computational time and complexity of the classifier. In this study, we use as secondary indexes:

- the ratio of the scalar products in band_i over band_j, $S_{ij} = S_{band_i} / S_{band_j}$,
- the product of the TFC based upon *SPWD* of band_i and band_j, $S_{ij} = S_{band_i} \cdot S_{band_j}$.

For simplicity, justification and definition of the selected bandwidth for this application are reported in the results section.

5.4. Results

The primary and secondary indexes previously described were computed in different frequency sub-bands. This gives a large number of parameters exceeding the number of subjects in the database. Thus the use of all these descriptors is not feasible.

Hence, a selection of the most relevant descriptors that help in the classification of patients according to their result in the HUTT is needed. Many sophisticated methods can be used to search the most effective features from the available features set. In this chapter, different combination of the feature selection algorithms (FS) and the classification methods described in chapter 3 are applied. For each combination, the classification procedure was repeated m times ($m=10$) taking randomly in each iteration 50% of the data for training and 50% for the testing phase. The average value of sensitivity (Se) and specificity (Sp) are computed to indicate the global performance of the classifier. This section is divided in four main sub-sections. The first one describes how the parameters were tuned, the second one proposes and describes the result of the classification procedure, and the last one proposes a comparison with the local linear correlation coefficient.

5.4.1. Frequency band selection and Optimization of SPWD parameters

The SPWD coherence method evaluating the correlation between signals was applied to study the dynamic interaction between:

- RR-interval time series and *AmpS* time series *RR/AmpS*;
- RR-interval time series and the *PTT* time series *RR/PTT*;
- RR-interval time series and the *dPdt_max* time series *RR/dPdt_max*.

The entire frequency range was divided into 10 sub-bands (figure 5.6), in each sub-band, primary and secondary parameters previously described were computed.

Bands limits	0-0.02	.02-0.2	0.2-0.4	0.4-0.6	0.6-0.8	0.8-1.0	1.0-1.2	1.2-1.4	1.4-1.6	1.6-1.8
Band number	1	2	3	4	5	6	7	8	9	10

FIG. 5.6 - The different sub-bands with their numbers as used in this paper.

Different combinations of temporal and spectral resolutions are obtained by changing the parameters of the kernel described in equation (5.7). The kernel which gave the smaller values of Δ_t and Δ_f and which respect the constraint $\gamma(t,f) \in [0,1]$ was chosen. In this paper, the kernel with $(\Delta_t, \Delta_f) = (10.78s, 0.038 \text{ Hz})$ was finally retained.

The distribution of time-frequency coherence $[\gamma^{band}(t,f)]_{th}$ was computed in the different sub-bands using several threshold values ($th=0.2; 0.3; \dots 0.9$) for each window. The threshold th which gives the most significant statistical difference (using the Mann-Whitney test) between the two groups was selected. After that, all the parameters $[\gamma^{band}(t,f)]_{th}$, secondary indexes of $[\gamma^{band}(t,f)]_{th}$, the scalar products S_{band} and secondary indexes of the scalar products, are the input of the feature selection and classification steps.

5.4.2. Classification performances

As mentioned in the previous section, different FS algorithms were evaluated based on different classifiers; all combinations (Relief-KSVM; Relief-KNN; SFS-KSVM; SFS-KNN; *Probe* feature-KSVM; *Probe* feature-KNN) gave approximately the same performance when trying minimizing approximately the error rate. Table 5.1 reports the best performance measured among all combinations in each condition for each time series couple. As mentioned before, they correspond to the mean value of specificity and sensitivity over the 10 test sets. The classification results presented in table 5.1 show that the performance of classification improves relatively for the different couples when the studied 5-min time interval is closer to the end of the test (Last5). However, the best accuracy is found in Early15 using only the scalar product of the HF band [0.2-0.4 Hz] of the *RR/AmpS* couple. It is worthwhile to mention that others couples in the same window present quiet good performance.

TAB. 5.1 - Optimal classification results obtained using parameters computed from *SPWD* coherence .

Condition	Time series	Sensitivity (Se)	Specificity (Sp)	Selected Parameters
Supine5	<i>RR/AmpS</i>	59.3%	58.6%	S_2/S_3
	<i>RR/dPdt_max</i>	64.6%	54.1%	$\gamma_4/\gamma_9; \gamma_5/\gamma_9$
	<i>RR/PTT</i>	62.5%	71.4%	S_2/S_3
Early5	<i>RR/AmpS</i>	72.1%	64.8%	$\gamma_6/\gamma_9; S_1/S_4; S_4/S_{10}; S_4/S_9; S_3/S_{10}; S_5/S_8$
	<i>RR/dPdt_max</i>	79.6%	53.4%	S_3/S_5
	<i>RR/PTT</i>	85.4%	41.7%	$\gamma_3 * \gamma_{10}; S_1; S_5/S_{10}; S_1/S_3; S_5/S_8$
Last5	<i>RR/AmpS</i>	68.2%	70.3%	$\gamma_3 * \gamma_6$
	<i>RR/dPdt_max</i>	61.1%	80%	$\gamma_1/\gamma_4; \gamma_9/\gamma_{10}; \gamma_1/\gamma_3; \gamma_1/\gamma_5$
	<i>RR/PTT</i>	77.9%	70.3%	$\gamma_2 * \gamma_7$
Early15	<i>RR/AmpS</i>	87.5%	98.3%	S_3
	<i>RR/dPdt_max</i>	68.9%	87.6%	$S_1/S_3; S_3/S_4; S_1/S_2$
	<i>RR/PTT</i>	70.7%	70%	$S_2/S_3; S_7/S_8; S_9/S_{10}; S_8/S_9; S_2 * S_3; \gamma_2/\gamma_6$

The best performance obtained during the Early15 with the *RR/AmpS* was derived when using the SFS-KNN while others combinations give approximately the same performance with different combination of feature selection and classification methods. The results are summarized in table 5.2.

TAB. 5.2 - Performance of classification during Early15

	Se	Sp	Selected parameters
Relief-KSVM	86.6%	98.4%	$S_4; S_3; S_9$
Relief-KNN	85.7%	98.6%	$S_4; S_3$
SFS-KSVM	70.4%	81.7%	$S_1/S_2; S_3; S_2/S_3; S_4/S_8; S_2^*S_5; \gamma_6/\gamma_8;$ $\gamma_2 * \gamma_4; \gamma_2 * \gamma_7; \gamma_1$
SFS-KNN	87.5%	98.3%	S_3
Probe feature-KNN	63.2%	73.8%	$\gamma_9/\gamma_{10}; S_4; S_3; S_8^*S_{10}$
Probe feature-KSVM	62.5%	94.5%	$\gamma_9/\gamma_{10}; S_4; S_3; S_8^*S_{10}; S_3^*S_4; \gamma_4$

In synthesis, the contingency matrix in table 5.3 represents the average result of the 10 iterations of the best performance when using 50% of population for the training process and the remaining part for the test.

TAB.5.3 - Matrix representing the average result of the 10 iterations of SFS-KNN during Early15 using S_3 on the test set.

	Classified as	
	Negative	Positive
Negative	13.9	0.1
Positive	1.7	12.3

We would like to mention here that in the previous chapter, the best performance based upon HRV were with sensibility equal to 88.5% and $Sp=80.6\%$, meaning that taking into account this inter-relationship improve the performance.

5.4.3. Comparison with local linear correlation coefficient

The local linear correlation coefficient $R^2(t,f)$ between the different couples described above was also computed using the method described in (Ansari-Asl et al., 2005) which estimates the local linear correlation coefficient by computing the Pearson Product-Moment correlation between two signals after applying a narrow-band filter bank. In others words, we computed:

$$R^2_{x,y}(t, f) = \max_{\tau_m \leq \tau \leq \tau_M} (r^2_{x,y}) \quad (5.5)$$

$$r^2_{x,y} = \frac{\left[\sum_{k=-\frac{H}{2}}^{\frac{H}{2}} x_f(k) y_f(k + \tau) \right]^2}{\sum_{k=-\frac{H}{2}}^{\frac{H}{2}} x^2_f(k) \sum_{k=-\frac{H}{2}}^{\frac{H}{2}} y^2_f(k + \tau)} \quad (5.6)$$

$r_{x,y}^2$ is the estimation of the linear correlation coefficient between the two zero-mean non stationary signals $x_f(t)$ and $y_f(t)$, outputs of a band-pass filter over a sliding window of duration H . The appropriate filter bank is developed from the Short-Time Fourier Transform (STFT). The parameters τ_m and τ_M are minimum and maximum bounds for the time delay between $x_f(t)$ and $y_f(t)$ respectively.

The same features, Distribution of time-frequency coherence over a threshold, were computed from the local linear correlation coefficient $R^2(t,f)$ to enter the feature selection algorithms. The selected parameters were also evaluated by the two classifiers KNN and KSVM. It is worthwhile mentioning that the scalar product parameters cannot be computed from the local linear correlation coefficient $R^2(t,f)$.

Table 5.4 summarizes the classification results obtained by applying this method using the same data mining procedure previously described. It showed that the use of local linear correlation coefficient in time-frequency domain provides lower performances than the *SPWD* coherence with a sensitivity and specificity that almost reach 75.2% and 66.8% during Early15. The results also indicate that applying this method, the *RR/PTT* is the most discriminate couple and Early15 is still the time interval that offers the best performance of classification.

TAB. 5.4 - Optimal classification results obtained using parameters computed from $R^2(t,f)$

Condition	Selected couple	Sensitivity (Se)	Specificity (Sp)	Parameters
Supine5	<i>RR/AmpS</i>	69.3%	54.5%	$R_1/R_6; R_1/R_3; R_5/R_{10}; R_1/R_7; R_2/R_7$
	<i>RR/ dPdt_max</i>	65%	56.2%	R_3/R_7
	<i>RR/ PTT</i>	48.2%	87.6%	$R_3; R_3^*R_7; R_7^*R_8;$
Early5	<i>RR/AmpS</i>	59.6%	66.9%	$R_6/R_{10}; R_4/R_5; R_4/R_{10}; R_1/R_2; R_1^*R_2;$ $R_4^*R_{10}$
	<i>RR/dPdt_max</i>	67.9%	51.7%	$R_4/R_8; R_3/R_8; R_2/R_3; R_4/R_7; R_3/R_7; R_1/R_2$
	<i>RR/PTT</i>	66.4%	52.7%	R_6/R_{10}
Last5	<i>RR/AmpS</i>	56.8%	66.5%	R_9/R_{10}
	<i>RR/dPdt_max</i>	67.07%	63.1%	$R_2^*R_{10}$
	<i>RR/PTT</i>	69.6%	50.7%	$R_5/R_{10}; R_2/R_3$
Early15	<i>RR/AmpS</i>	58.9%	72.7%	$R_3; R_2/R_3$
	<i>RR/dPdt_max</i>	65.4%	68.6%	$R_2/R_3; R_1^*R_8; R_1^*R_2$
	<i>RR/PTT</i>	66.8%	75.2%	$R_5/R_{10}; R_2/R_3; R_1^*R_7; R_3^*R_7$

5.5. Discussion

Although, it is clear that a disturbance in the autonomic nervous system is the origin of vasovagal syncope (Hayoz et al., 1996) (Thomson et al., 1997), the mechanism

of this syncope is complex and still incompletely understood. Several studies have reported approaches using HRV (Pruvot et al., 1994) (Kochiadakis et al., 1997) or blood pressure variability (Pitzalis et al., 2002) alone. However, vasovagal syncope during HUTT is an abnormal response of the baroreflex, which includes cardiac and vascular components. Thus, the analysis of a single measurement can be insufficient to reflect the underlying physiological processes. Therefore, analyzing parameters that characterize the interaction between heart rate and blood pressure can be of interest.

This chapter presents a time-frequency method for characterizing the dynamic interaction between different cardiovascular signals, to predict the outcome of HUTT before its ending. The data have been recorded from 57 subjects undergoing HUTT, of which 28 have experienced syncope during the test.

Using *SPWD* coherence in *Supine5* can lead to a performance with 62.5% of sensitivity and 71.4% of specificity using parameters extracted from *RR/PTT*. This performance increases in *Early5* using the *RR/AmpS* couple (Se=72.1%; Sp=64.8%). However the results are more discriminative in *Last5* than the other 5-min windows (Se=77.9%; Sp=70.3%) using *RR/PTT*. The best accuracy of the classifier is found in *Early15*, using the scalar product (S_3) parameter of TF spectrum of *RR* and *AmpS* time series in bandwidth [0.2-0.4] Hz (Se=87.5%; Sp=98.3%). This demonstrates that using 15-min window in the tilting position can lead to early distinguishing negative and positive subjects. Indeed the simple scalar product of the auto spectrum of *RR* and *AmpS* in the sub-band 0.2-0.4 Hz differs significantly between both groups. The between-group comparison demonstrates that S_3 is significantly lower in non-fainter than in fainter group (p -value<0.01). This suggests that the similarity between *RR*-interval and *AmpS* time series in HF band is higher in fainters than in non-fainters during the first 15 min of tilt test. Typical auto TF spectrum for *RR*-interval and *AmpS* time series extracted from a negative and a positive subject during *Early15* are shown in figure 5.7.

Our findings are complementary to the results published in (M. Orini et al., 2012), where they noticed an abrupt reduction in the cardiovascular coupling in the HF band in negative subjects. The absence of the reduction in similarities between cardiovascular signals in positive subjects reported in our study may play a role in the occurrence of the syncope.

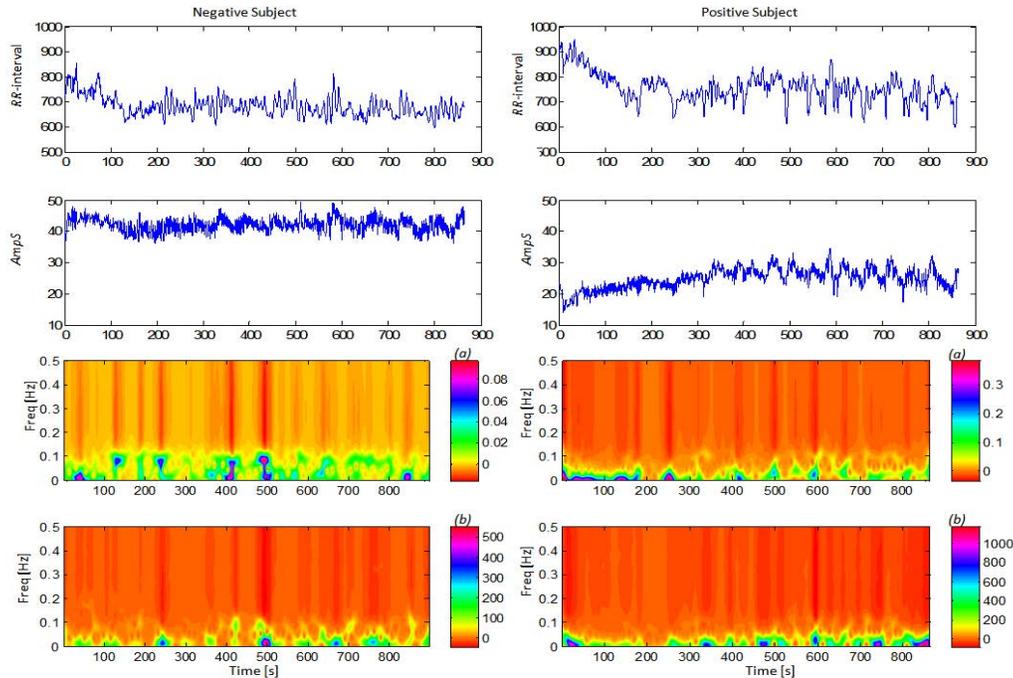


FIG. 5.7- Representation of the *RR*-interval time series and the *AmpS* time series and their corresponding TF auto spectrum respectively (a) and (b) for a negative (non-fainters) and a positive (fainters) subject during Early15.

5.6. Conclusion

This study aimed to characterize the time-frequency interactions between cardiovascular signals using the *SPWD* and to evaluate its ability to predict the outcome of HUTT, *i.e.* if subject will faint or not. For this purpose several new indexes were proposed to quantify the relationship between signals: the scalar product in specific bandwidth, the ratio and product of two scalar products, and the distribution of time-frequency coherence over a threshold.

Our results demonstrate for the first time that studying the dynamic interaction between *RR*-interval and the amplitude of systolic blood pressure during the first 15 minutes of HUTT appears very interesting as it provides prominence results ($Se=87.5\%$; $Sp=98.3\%$) in predicting the outcome of this test. This performance is higher than other performances computed using the relationship between *RR*-interval and *PTT* or *RR*-interval and *dPdt_{max}* time series.

During the first 15 minutes of the HUTT, this study also shows that in high frequency band [0.2-0.4Hz], the similarity between the *RR*-interval and the *AmpS* time series (reflected by the scalar product of the auto time frequency spectrum computed with *SPWD* of these two signals) is higher in positive than in negative subjects.

From a methodological point of view, a comparison was carried out between the proposed *SPWD* method and the linear correlation coefficient in the time and frequency domains as suggested in (Ansari-Asl et al., 2005). This comparison highlights the powerfulness of the *SPWD* method in predicting the syncope during HUTT (Another comparison study with an approach based on partial least square regression is voluntarily reported in appendix B And confirms these interesting results). In the same manner, comparison between several feature selections methods (Relief method, sequential forward selection, *probe* feature method) and classification approaches (SVM and KNN) do not clearly exhibit the significant powerfulness of one combination compared to other for our purpose.

Transposed to the clinical context of syncope diagnosis, these results suggest that the *SPWD* may be a very useful analyzing technique to predict the outcome of the HUTT in the early 15 min of the test. This leads to a significant reduction of HUTT duration which is usually from 30 to 45 minutes (Voice et al., 1998). Thus, *SPWD* could lead to minimizing the duration of an uncomfortable diagnosis procedure, and to improve the efficiency of syncope management units in hospitals.

Appendix B

Data analysis with partial least square regression

XLSTAT-PLS is a complementary statistical module for Microsoft Excel that contains advanced analytical functions such as Partial Least Squares regression (PLSR) or principal component regression (PCR). These regression methods are used to overcome some limitations of the traditional linear regression and analysis of variance as the non-collinearity of the explanatory variables and the minimum number of samples required. These properties make *XLSTAT-PLS* software an excellent complement to the analysis of large data sets (Huffel, 1997). In this study we decided to use the PLSR to analyze our data during the time-interval that gives the best classification performance (Early15) and to verify the importance of the parameter S_3 . The PLSR is a fast, efficient and optimal method for a minimization criterion of covariance and it is recommended when a large number of explanatory variables is used.

Principle of PLSR

The idea of the PLS (Partial Least Squares) is to create from an array of n observations described by p variables, a set of h components with $h < p$. The method of construction of the components differs from that of the principal component analysis PCA, and has the advantage of well accommodate the presence of missing data. The number of components to retain is usually determined based on a criterion involving a cross validation.

Graphical results of PLSR

PLSR was applied on a learn dataset composed from half the population using only primary parameters then the algorithm is tested on the rest of the population. The chart represented in figure B1 shows the quality of PLSR based on the number of selected components. The Q^2 cumulated index measures the global goodness of fit and the predictive quality of the models composed by the variables. *XLSTAT-PLS* has automatically selected 3 components. The cumulated R^2Y and R^2X cum that correspond to the correlations between the explanatory (X) and dependent (Y) variables with the components are close to 1 with components. This indicates that the components generated by the Partial Least Squares regression summarize both the X s and the Y s. Figure B2 displays the VIPs (Variable Importance for the Projection) for each

explanatory variable, for an increasing number of components. This allows to quickly identifying which are the explanatory variables that contribute the most to the models and this verifies the importance of the parameter S_3 . Figure B3 illustrates the ROC curve of the testing data.

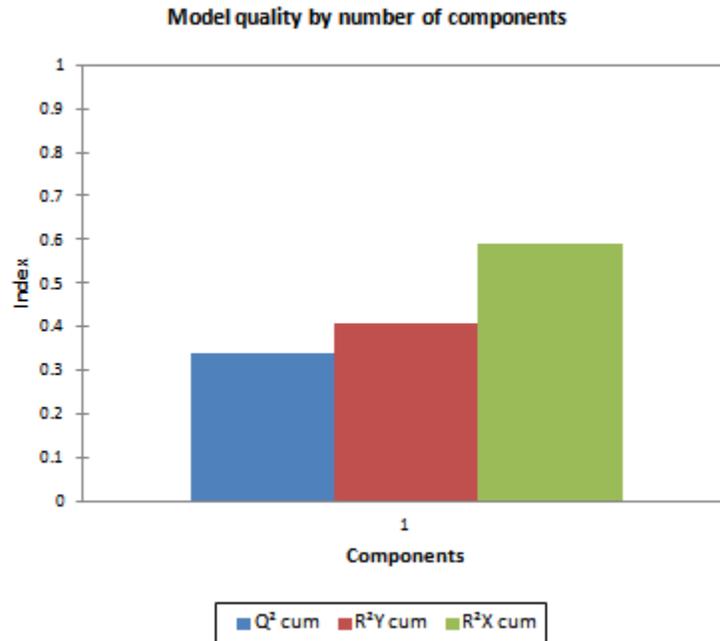


FIG. B1 - Chart representing mode quality by number of components.

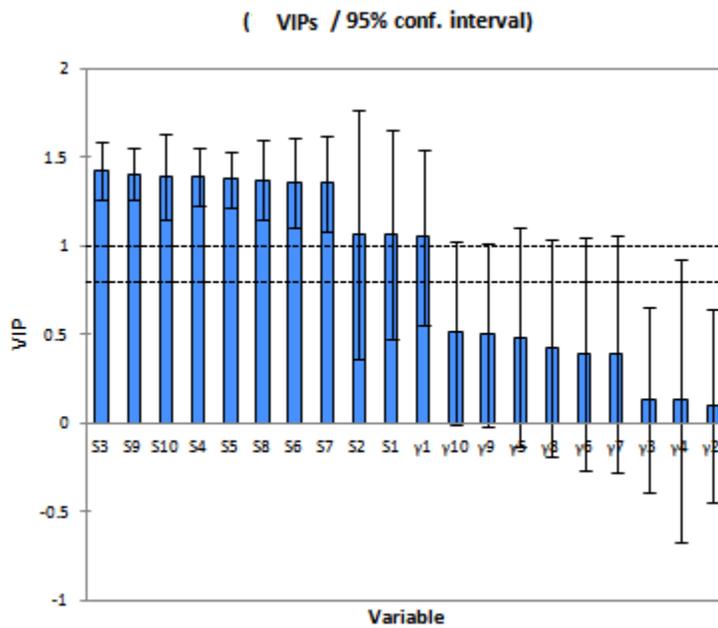


FIG. B2 - Variable Importance for the Projection

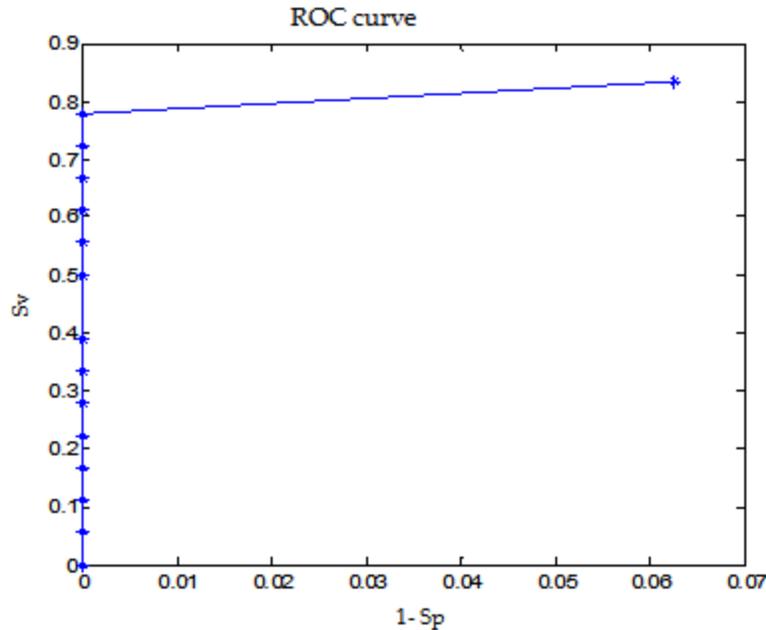


FIG . B3 - ROC curve of the testing phase

References

- Ansari-Asl, K., Bellanger, J.-J., Bartolomei, F., Wendling, F., Senhadji, L., 2005. Time-frequency characterization of interdependencies in nonstationary signals: application to epileptic EEG. *IEEE Trans. Biomed. Eng.* 52, 1218–1226. doi:10.1109/TBME.2005.847541
- Bellard, E., Fortrat, J.-O., Vielle, B., Dupuis, J.-M., Victor, J., Lefthériotis, G., 2001. Early predictive indexes of head-up tilt table testing outcomes utilizing heart rate and arterial pressure changes. *Am. J. Cardiol.* 88, 903–906. doi:10.1016/S0002-9149(01)01904-X
- Carrault, G., Beuchee, A., Pladys, P., Senhadji, L., Hernandez, A., 2009. Time-frequency relationships between heart rate and respiration: A diagnosis tool for late onset sepsis in sick premature infants, in: *Computers in Cardiology, 2009*. Presented at the *Computers in Cardiology, 2009*, pp. 369–372.
- Dreyfus, G., Martinez, J.-M., Samuelides, M., Gordon, M.B., Badran, F., Thiria, S., 2011. *Apprentissage statistique: Réseaux de neurones - Cartes topologiques - Machines à vecteurs supports*. Editions Eyrolles.
- Dumont, J., Hernandez, A.I., Carrault, G., 2010. Improving ECG Beats Delineation With an Evolutionary Optimization Process. *IEEE Trans. Biomed. Eng.* 57, 607 –615. doi:10.1109/TBME.2008.2002157
- Gimeno-Blanes, F.J., Rojo-Álvarez, J.L., Caamaño, A.J., Flores-Yepes, J.A., García-Alberola, A., 2011. On the feasibility of tilt test outcome early prediction using

- ECG and pressure parameters. *EURASIP J. Adv. Signal Process.* 2011, 33. doi:10.1186/1687-6180-2011-33
- Hayoz, D., Noll, G., Passino, C., Weber, R., Wenzel, R., Bernardi, L., 1996. Progressive withdrawal of muscle nerve sympathetic activity preceding vaso-vagal syncope during lower-body negative pressure. *Clin. Sci. Lond. Engl.* 1979 91 Suppl, 50–51.
- Huffel, S. van, 1997. *Recent Advances in Total Least Squares Techniques and Errors-in-variables Modeling.* SIAM.
- Karapetian, G.K., Engels, H.J., Gretebeck, K.A., Gretebeck, R.J., 2012. Effect of caffeine on LT, VT and HRVT. *Int. J. Sports Med.* 33, 507–513. doi:10.1055/s-0032-1301904
- Khodor, N., Matelot, D., Carrault, G., Amoud, H., Khalil, M., Ville, N., Carre, F., Hernandez, A., 2014. kernel based support vector machine for the early detection of syncope during head-up tilt test. *Physiol. Meas.*
- Kochiadakis, G.E., Orfanakis, A., Chrysostomakis, S.I., Manios, E.G., Kounali, D.K., Vardas, P.E., 1997. Autonomic nervous system activity during tilt testing in syncopal patients, estimated by power spectral analysis of heart rate variability. *Pacing Clin. Electrophysiol. PACE* 20, 1332–1341.
- Kouakam, C., Lacroix, D., Zghal, N., Logier, R., Klug, D., Le Franc, P., Jarwe, M., Kacet, S., 1999. Inadequate sympathovagal balance in response to orthostatism in patients with unexplained syncope and a positive head up tilt test. *Heart* 82, 312–318.
- Lippman, N., Stein, K.M., Lerman, B.B., 1995. Failure to decrease parasympathetic tone during upright tilt predicts a positive tilt-table test. *Am. J. Cardiol.* 75, 591–595.
- Madrid, A.H., Moro, C., Marín-Huerta, E., Novo, L., Mestre, J.L., Lage, J., Ricoy, E., 1994. [Usefulness of the RR variability in the diagnosis of neurogenic syncope]. *Rev. Esp. Cardiol.* 47, 536–543.
- Mainardi, L.T., Bianchi, A.M., Cerutti, S., 2002. Time-frequency and time-varying analysis for assessing the dynamic responses of cardiovascular control. *Crit. Rev. Biomed. Eng.* 30, 175–217.
- Mallat, Z., Vicaut, E., Sangaré, A., Verschueren, J., Fontaine, G., Frank, R., 1997. Prediction of head-up tilt test result by analysis of early heart rate variations. *Circulation* 96, 581–584.
- Matz, G., Hlawatsch, F., 2000. Time-frequency coherence analysis of nonstationary random processes, in: *Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing, 2000.* Presented at the Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing, 2000, pp. 554–558. doi:10.1109/SSAP.2000.870186
- Movahed, M.R., Hassapoyannes, C.A., 2001. Prediction of non-occurrence of syncope during a tilt-table test by early heart rate variations. *J. S. C. Med. Assoc.* 1975 97, 207–210.

- Orini, M., Bailon, R., Mainardi, L.T., Laguna, P., Flandrin, P., 2012. Characterization of Dynamic Interactions Between Cardiovascular Signals by Time-Frequency Coherence. *IEEE Trans. Biomed. Eng.* 59, 663–673. doi:10.1109/TBME.2011.2171959
- Orini, M., Laguna, P., Mainardi, L.T., Bailón, R., 2012. Assessment of the dynamic interactions between heart rate and arterial pressure by the cross time-frequency analysis. *Physiol. Meas.* 33, 315–331. doi:10.1088/0967-3334/33/3/315
- Pitzalis, M., Massari, F., Guida, P., Iacoviello, M., Mastropasqua, F., Rizzon, B., Forleo, C., Rizzon, P., 2002. Shortened head-up tilting test guided by systolic pressure reductions in neurocardiogenic syncope. *Circulation* 105, 146–148.
- Pruvot, E., Vesin, J.M., Schlaepfer, J., Fromer, M., Kappenberger, L., 1994. Autonomic imbalance assessed by heart rate variability analysis in vasovagal syncope. *Pacing Clin. Electrophysiol. PACE* 17, 2201–2206.
- Schang, D., Feuilloy, M., Plantier, G., Fortrat, J.-O., Nicolas, P., 2007. Early prediction of unexplained syncope by support vector machines. *Physiol. Meas.* 28, 185. doi:10.1088/0967-3334/28/2/007
- Sumiyoshi, M., Nakata, Y., Mineda, Y., Tokano, T., Yasuda, M., Nakazato, Y., Yamaguchi, H., 2000. Does an early increase in heart rate during tilting predict the results of passive tilt testing? *Pacing Clin. Electrophysiol. PACE* 23, 2046–2051.
- Sutton, R., Brignole, M., Benditt, D.G., 2012. Key challenges in the current management of syncope. *Nat. Rev. Cardiol.* 9, 590–598. doi:10.1038/nrcardio.2012.102
- Thomson, H.L., Wright, K., Frenneaux, M., 1997. Baroreflex Sensitivity in Patients With Vasovagal Syncope. *Circulation* 95, 395–400. doi:10.1161/01.CIR.95.2.395
- Virag, N., Sutton, R., Vetter, R., Markowitz, T., Erickson, M., 2007. Prediction of vasovagal syncope from heart rate and blood pressure trend and variability: experience in 1,155 patients. *Heart Rhythm Off. J. Heart Rhythm Soc.* 4, 1375–1382. doi:10.1016/j.hrthm.2007.07.018
- Voice, R.A., Lurie, K.G., Sakaguchi, S., Rector, T.S., Benditt, D.G., 1998. Comparison of tilt angles and provocative agents (edrophonium and isoproterenol) to improve head-upright tilt-table testing. *Am. J. Cardiol.* 81, 346–351.
- White, L.B., Boashash, B., 1990. Cross spectral analysis of nonstationary processes. *IEEE Trans. Inf. Theory* 36, 830–835. doi:10.1109/18.53742

Chapter 6

Evaluation of syncope detection coupling a novel phase space analysis algorithm and Hidden Semi-Markov models

In this chapter, we evaluate the ability to automatically predict syncope and differentiate between patients with negative and positive response to HUTT, using only 12-min RR-interval time series for the supine position, preceding HUTT. The proposed method is unlike the traditional approaches, not based on instantaneous measures but relying on the analysis of the dynamics of the extracted time series. An original method - based on the analysis of cardiovascular dynamics, a combination of phase space (PS) analysis and kernel support vector machine (KSVM)-, is proposed. The dynamic behavior of the RR-interval time series was first analyzed using reconstructed phase space (RPS). Parameters computed from the phase space area such as the phase space density and indices derived from the recurrence quantification analysis were computed. Only parameters, displaying a statistical difference, are used for further classification using KSVM, to identify negative and positive patients. RPSs of time series extracted from other cardiovascular signals in different time windows were also constructed and the aforementioned parameters are computed to compare their capability to predict the outcome of the test with those obtained in 12 min of supine position.

In the second part, we examine the capability of a new method to predict syncope. The extracted time series and their dynamics are here characterized by a set of Hidden Semi-Markov Models (HSMM). The prediction of the syncope is based on the comparison of the likelihood that a given time series has generated by a negative HSMM (learned using time series from negative subjects) or positive HSMM (learned using time series from positive subjects).

6.1. Analysis of the dynamics of time series

The characterization of the dynamics of time series is a challenging problem. Two main methodological approaches to this problem can be distinguished: i) those based on pattern recognition theory, associated with a similarity measure (such as the Dynamic Time Warping distance) either computed directly from the raw data, or from a transformation of these data into a specific feature space and ii) those based on model estimation. The former approach requires little prior knowledge and is well suited to simple classification and detection problems. However, it may be difficult to apply to biomedical data, due to the significant intra and inter-patient variability and the heterogeneity of the dynamics in the data. The latter approach allows a direct representation of multidimensional time series and the introduction of a priori knowledge can be achieved at several levels of detail, depending on the intended application. Model-based approaches can be used to quantify the likelihood that a given observation is generated by a given model (classification problem) or to estimate a future state of the system, through simulation. However, these approaches are generally more difficult to develop and are particularly dependent on the quantity and quality of the data used in the learning and validation phases of the model.

In this chapter, we have chosen a model-based approach. This choice is justified by the fact that i) the result of the HUTT can be used as a source of a priori knowledge for this application, ii) several short-time windows may contain useful information on the dynamics of the times series (for example, during the early beginning of the tilt) and these windows must be analyzed in a joint manner (multivariate analysis) and iii) the dynamics of the different time-windows studied are heterogeneous, because they are the result of complex physiological processes operating with very different time constants. A Brief state of the art on time series modeling is described below:

- a) *Kalman filters*: Kalman filters have been widely used in the fields of linear and non-linear system analysis, or trajectory tracking (Cassidy and Penny, 2002). They have also been applied successfully to the biomedical field (Gordon and Smith, 1990). In particular, the approach provides a generic framework to compute probabilities of whether changes have occurred as well as to identify the nature of these changes. A recent extension of particle filters to online non-linear/non gaussian bayesian tracking represents an elegant alternative. However, the design of the particle filter for a given application remains critical due to its sensitivity to the choice of the density law. In this sense, the application of Kalman filters and its extensions remains limited and it is thus inappropriate in our case.
- b) *Artificial neural network models*: Some specific artificial neural network (ANN)

architectures are particularly suited to the characterization of time series (Time Delay Neural Networks (Waibel et al., 1989), Recurrent Neural Networks (J. Connor, 1991) and have been used on many applications, including the biomedical field (Leistritz et al., 2006), (Vásquez et al., 2001). However, the optimal structure of these ANN depends on a number of parameters that can be difficult to define (number of layers, neurons in each layer, etc...). Also, the black box nature of this method avoids the possibility of interpreting the network parameters obtained after training.

- c) *Trajectories embedded in phase spaces*: The reconstructed phase space (RPS) is one of the well-known techniques that can be applied to analyze the dynamics of non linear systems, by transforming time-lagged copies of the initial signal onto axes of a new high dimensional space. The use of RPSs has provided interesting results in different applications, such as in speech synthesis and recognition (Kumar and Mullick, 1996) (Petry and Barone, 2002) as well as in the biomedical field (Roberts et al., 2001).
- d) *Hidden Markov models (HMM)*: HMM have been widely and successfully applied in the field of speech recognition (Yamagishi et al., 2006) and also in the biomedical context (Faisan et al., 2005). Extensions of the original HMM have been proposed in order to represent continuous observations, through continuous density HMM (CDHMM), or to better represent the temporal dynamics of state transitions (Hidden Semi-Markov models-HSMM). Recent improvements such as hybrid Markov models, combining, for instance, an HMM and an Artificial Neural Network (Tang, 2009), underline the generic background and flexibility of this theory.

According to the preceding discussion, Reconstructed Phase Space and Hidden Semi-Markov Models represent an interesting approach for characterizing the dynamics of time series and can be useful tools to analyze the complex cardiovascular system. For these reasons, we propose to evaluate the capability of these two approaches to figure out our problem. So, this chapter is divided into two parts:

- In the first one, we exploit the RPS approach, thus we first define the principle of the phase space algorithm and the integration of a time series in a phase space, and then we describe the most practical methods used to choose the time delay and the embedding dimension. After that, we present the parameters extracted from the phase space. In the results section, we display the outcomes of the classification by KSVM using the RR time series of 12 min of supine position; these results are then verified using the partial least square regression and compared with other performance obtained by time series extracted from the blood pressure signal in

different time windows. Finally, an optimal fusion decision rule (Hernandez et al., 1999) was used to evaluate the performance of the different time series combined together.

- The second part is dedicated to test the HSMM approach, so the HSMM architecture and the algorithms used to estimate the model parameters are presented. The syncope prediction process is described and then evaluated, to end up with a conclusion.

6.2. Reconstructed phase space

6.2.1. Definition

The phase Space is defined as a representation space used mainly in mathematics and physics to demonstrate and visualize the changes in the dynamic variables of a system. Each of these variables is represented by an axis of the phase space and for each possible state of the system a point is then drawn in this space. The succession of these points represents the dynamic evolution of the system and the resulting shape allows to exhibit properties of the system that are not visible otherwise (attractors, fixed points, etc.). Moreover, the space formed in this path is constant over time; thus it is possible to calculate the state of the system in the future (or in the past) by using the transfer function f :

$$X(t+1)=f(X(t)) \quad (6.1)$$

Where $X(t)$ denotes an observable variable of the system at time t . Finally, as each point of the phase space has a temporal successor, two different paths may not intersect.

When the observed variables are analyzed and not the state variables, a reconstructed phase space (RPS, also called pseudo-phase space) is created. In this case, the axes incorporate different delayed versions of the same variable:

$$[x(t-\tau), x(t-2\tau), \dots, x(t-M\tau)] \quad (6.2)$$

In theory, Takens theorem (Takens, 1981) specifies that by correctly choosing these lagged variables (τ, M) and in case of non-noisy observations of deterministic systems, integrated RPS trajectories have, in a topological point of view, the same properties of real attractors, as shown in the space phases.

It is interesting to note here that the Poincaré plots, often used in heart rate variability analysis, are indeed RPS with particular values of M and τ .

6.2.2. Integration of time series in RPS

The creation of RPS requires the determination of the time delay τ and the embedding dimension M of the phase space. So, given an observation represented in time series $x(t_i)$ with $i=1, \dots, T$ (T is the number of points); the vector in reconstructed phase space at time t_i can be written as following:

$$X(t_i)=[x(t_i), x(t_i+\tau), x(t_i+2\tau), \dots, x(t_i+(M-1)\tau)], \quad (6.3)$$

In the theoretical case of non-noisy and infinite time series, Takens demonstrates that a dimension greater than $2 \times$ the fractal dimension provides a path in the pseudo phase space with an equivalent topology to that of the initial phase space, regardless the value of τ .

Out of this theoretical framework, Buzug (Buzug and Pfister, 1992) or Cellucci (Cellucci et al., 2003) reveal the importance of the choice of M and τ . For example, a very small value of time delay will give highly correlated points between the different axes and the really important correlations will not be visible, while too large time delay value will give points that seem randomly distributed in the RPS. For the embedding dimension M , an overestimation will complicate in a useless way the exploitation of the trajectory while an underestimate will create a folded trajectory. In this latter case, a position at time t does not necessarily give a unique position at time $t+1$. The minimization of mutual information is frequently used to choose the time delay and gives satisfactory results (Fraser and Swinney, 1986) (Kraskov et al., 2004) while the algorithm of the false nearest neighbors (Kennel et al., 1992) estimates iteratively the sufficient dimension to avoid non-significant trajectory folding.

6.2.2.1. Estimation of the time delay

To determine a suitable time delay, we used average mutual information (AMI). This method was proposed by Fraser and Swinney (Fraser and Swinney, 1986). The idea is that a good choice for τ is the one that, given $X(t)$, provides new information with $X(t+\tau)$ measurement. Given two time series U and V , the average mutual information between these two time series is defined by (Abarbanel, 1997):

$$I_{UV}(u_i, v_j) = \sum_{u_i, v_j} P_{UV}(u_i, v_j) \log_2 \left[\frac{P_{UV}(u_i, v_j)}{P_U(u_i)P_V(v_j)} \right] \quad (6.4)$$

Where: $P_U(u_i)$ is the probability of occurrence of u_i in set U .

$P_V(v_i)$ is the probability of occurrence of v_i in set V .

$P_{UV}(u_i, v_i)$ is the probability of co-occurrence of u_i in set U and v_i in set V .

Replacing U by $x(t_1), x(t_2), \dots, x(t_i)$ and V by $x(t_1 + \tau), x(t_2 + \tau) \dots x(t_i + \tau)$ then the average mutual information that we can estimate x in time $t + \tau$ from x in time t , $I(t)$ becomes:

$$I(\tau) = \sum_{t_i, t_i + \tau} P(x(t_i), x(t_i + \tau)) \log_2 \left[\frac{P(x(t_i), x(t_i + \tau))}{P(x(t_i)) P(x(t_i + \tau))} \right] \quad (6.5)$$

According to (Fraser and Swinney, 1986) the suitable value of τ is corresponds to the first local minimum of $I(\tau)$.

6.2.2.2. Choosing the embedding dimension

The false nearest neighbors (FNN) is a well known method used to estimate the embedding dimension of RPS (Kennel et al., 1992). The basic idea is, if the trajectory is projected into a space of too small dimension, which is insufficient to unfold the trajectory, false nearest neighbors appear. By increasing the dimension, the number of false nearest neighbors drops to zero and trajectory self-crossings decrease. (Figure 6.1) is an example taken from (Kennel et al., 1992) that shows trajectory with $M=1$ and $M=2$ for the x coordinates of a Hénon map. It demonstrates that the points A and B are false neighbors whereas A and C are real neighbors.

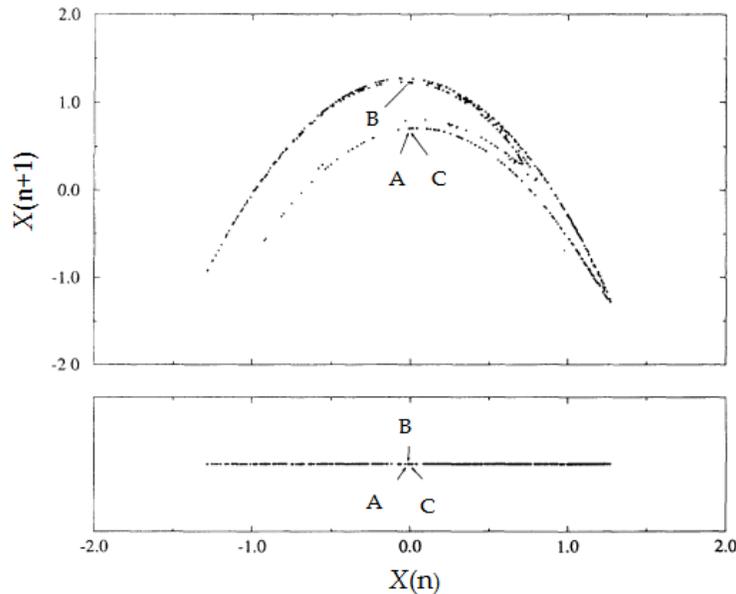


FIG. 6.1 - An example of trajectory with $M=1$ and $M=2$ for the x coordinates of a Hénon map.

Denoting by $X^N(t_i)$ the neighbor of the point $X(t_i)$ in the dimension M and referring to equation (6.3), the square of the Euclidean distance between these two points is (Kennel et al., 1992):

$$R_M^2(t_i) = \sum_{n=0}^{M-1} [x(t_i + n\tau) - x^N(t_i + n\tau)]^2 \quad (6.6)$$

Now passing to a higher dimension $M+1$, an $x(t_i + M\tau)$ is added onto $X(t_i)$, and the distance between $X(t_i)$ and $X^N(t_i)$ becomes:

$$R_{M+1}^2(t_i) = R_M^2(t_i) + [x(t_i + M\tau) - x^N(t_i + M\tau)]^2 \quad (6.7)$$

We denote $X^N(t_i)$ as a false neighbor of the point $X(t_i)$, if when passing from dimension M to dimension $M+1$, the distance between $X^N(t_i)$ and $X(t_i)$ increases, so the ratio $R_a(t_i)$ is equal or higher than a threshold R_{th} :

$$R_a(t_i) = \frac{|x(t_i + M\tau) - x^N(t_i + M\tau)|}{R_M(t_i)} \geq R_{th} \quad (6.8)$$

We also used a second criterion, combined with the first one to test the falsity of the neighbors:

$$\frac{|x(t_i + M\tau) - x^N(t_i + M\tau)|}{R_A} \geq A_\tau \quad (6.9)$$

A_τ is around 2 (Abarbanel, 1997) with R_A is the standard deviation of x .

6.2.3. The Phase space density

After the estimation of the time delay τ and the embedding dimension M by the two aforementioned methods, the corresponding phase space plot is scaled, to have all values within the range $[0,1]$ on each axis of the PS. After that, the scaled volume is divided into 1000 equally-sized cells (Srinivasan et al., 2003), each cell is referenced by an attributed number, according to its bounds in the axes (see algorithm 1). Table 6.1 shows an example of the first 3 cells of a decomposed 3 dimensions phase space. Then, the distribution of points in each cell is estimated. This gives an idea of the trend and the distribution of the trajectory in the phase space.

TAB. 6.1 - Example of the first 3 cells of a decomposed phase space.

Cell number	Bounds in 1st axis	Bounds in 2nd axis	Bounds in 3rd axis
1	0-0.1	0-0.1	0-0.1
2	0-0.1	0-0.1	0.1-0.2
3	0-0.1	0-0.1	0.1-0.3

Algorithm 1 : Decomposition of phase space area into 1000 cells

INITIALIZATION

/* Matrix containing lower and upper bounds of each cell in each axis*/

X,Y,Z;

/*Vector containing the number of each cell*/

NB=0;

LOOP FOR DECOMPOSITION

For (ix:=0) **while** (ix<0,9) **do** /* Decomposition of X axis*/

For (iy:=0) **while** (iy<0,9) **do** /* Decomposition of Y axis */

For (iz:=0) **while** (iz<0,9) **do** /* Decomposition of Z axis */

 X=[X ; ix ix+0,1];

 Y=[Y ; iy iy+0,1];

 Z=[Z ; iz iz+0,1];

 NB=[NB; NB+1];

 ix=ix+0,1;

 iy=iy+0,1;

 iz=iz+0,1;

EndFor

End

6.2.4. Recurrence quantification analysis

6.2.4.1. Recurrence plot

The recurrence of states means that states are arbitrary near to each other after some time. This is a fundamental property of deterministic dynamical systems and it has been known and discussed in several publications (Argyris et al., 1994) (Ott, 2002) (Monk and Compton, 1939). In 1987, Eckmann et al. (Eckmann et al., 1987) have proposed a technique to visualize the recurrence of states in the phase space. In fact, the dimension of PS can be higher than three; in this case it cannot be plotted only if it is projected into the two or three dimensional sub-spaces. The Eckmann's method permits the representation of the M -dimensional phase space trajectory through a 2-dimensional illustration of its recurrences: this illustration is called recurrence plot (RP).

The recurrence plot (RP) is an advanced method to analyze non-linear data. It is a representation of the phase space trajectory recurrence to a certain state, which shows the times when the trajectory of the PS passes approximately in the same area. The RP consists of a square matrix (figure 6.2), with elements corresponding to the times at which a state at time t_i recurs at a different time t_j (columns and rows correspond to a certain pair of times, *i.e.* t_i and t_j). This is marked as black and white dots in a plot (ones and zeros respectively in the matrix), this plot having thus two time axes. The matrix of RP can be computed by the following mathematical expression:

$$R_{i,j} = \Theta(\varepsilon - \|X(t_i) - X(t_j)\|); \quad i, j=1, \dots, N \quad (6.10)$$

with N being the number of states, ε is a threshold distance, $\|\cdot\|$ is a norm such as the Euclidean norm and $\Theta(\cdot)$ is the Heaviside function. The threshold ε is a constant which delineates a sphere centered at $X(t_i)$, if $X(t_j)$ belongs to this sphere so it will be close to $X(t_i)$. In this chapter ε is chosen according to literature (Eckmann et al., 1987) equal to 0.2 and the Euclidean distance is used. The Heaviside's step function is:

$$\Theta(\alpha) = \begin{cases} 0 & \text{if } \alpha < 0 \\ 1 & \text{if } \alpha \geq 0 \end{cases} \quad (6.11)$$

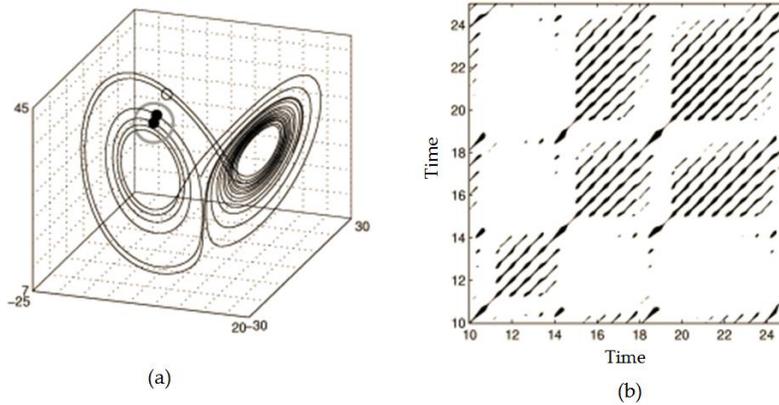


FIG. 6.2 - A typical example of a Lorenz system (a) and its recurrence plot (b) (Roberto Barbera, 2011).

As shown in figure 6.2, the RP displays characteristics caused by the dynamical behavior of the system (Eckmann et al., 1987) (Marwan et al., 2007), such as:

- Diagonal line: it corresponds to epochs where segments of phase space trajectory run parallel, the length of this diagonal line is determined by the duration of such similar local evolution of this trajectory segment.
- Vertical line: designates a time length in which a state does not change or changes slowly, it is an indication of laminar states.

6.2.4.2. Quantitative analysis of the RP

The recurrence quantification analysis (RQA) is a method to quantify the number and duration of state recurrences of a trajectory in a phase space. The main advantage of RQA is that it provides useful information, even for short and non-stationary data. RQA was developed first by Zbilut and Webber (Zbilut and Webber, 1992) which define measures using the recurrence point density and the diagonal structures in RP. After that, Gao (Gao and Cai, 2000) has introduced a recurrence time

statistics that corresponds to vertical structures. Recently, Marwan et al. have extended new measures of complexity based on diagonal and vertical structures in RP, the measures are summarized in table 6.2 (Marwan et al., 2007).

TAB. 6.2 - RQA parameters (Marwan et al., 2007)

Parameter	Definition
Recurrence Rate $R_e R_a$	The percentage of recurrence points in an RP, it corresponds to the correlation sum: $R_e R_a = \frac{1}{N^2} \sum_{i,j=1}^N R_{i,j}$
Determinism DET	The percentage of recurrence points which form diagonal lines: $DET = \frac{\sum_{l=l \min}^N l P(l)}{\sum_{i,j} R_{i,j}}$ <p>$P(l)$ is the histogram of the lengths l of the diagonal lines.</p>
Laminarity LAM	The percentage of recurrence points which form vertical lines: $LAM = \frac{\sum_{v=v \min}^N v P(v)}{\sum_{v=1}^N v P(v)}$ <p>$P(v)$ is the histogram of the lengths v of the vertical lines</p>
Average diagonal line length $\langle L \rangle$	The average length of the diagonal lines: $L = \frac{\sum_{l=l \min}^N l P(l)}{\sum_{l=l \min}^N P(l)}$
Trapping Time TT	The average length of the vertical lines: $TT = \frac{\sum_{v=v \min}^N v P(v)}{\sum_{v=v \min}^N P(v)}$
Longest diagonal line L_{max}	The length of the longest diagonal line: $L_{max} = \max(\{ l_i ; i = 1 \dots N_l \})$ <p>N_l: number of diagonal lines in the RP</p>
Longest vertical line V_{max}	The length of the longest vertical line: $V_{max} = \max(\{ v_i ; i = 1 \dots N_v \})$ <p>N_v: number of vertical line in Rp</p>
Entropy $ENTR$	The Shannon entropy of the probability distribution of the diagonal line lengths $p(l)$:

	$ENTR = - \sum_{l=l \min}^N p(l) \ln p(l)$
Recurrence period density entropy <i>RPDE</i>	$RPDE = -(\ln L_{\max})^{-1} \sum_{l=l \min}^N p(l) \ln p(l)$

6.2.5. Results

Subjects enrolled in the study are 66 men described in previous chapters, 31 patients had a negative response to HUTT and 35 patients have developed syncope during the test (positive response). The next subsections represent the adjustment of the embedding dimension and delay of the RPS, and the classification performance in the 12 min of supine position using RR-interval time series. The performance using partial least square regression method is also evaluated. In order to compare the results of the 12 min of RR-interval time series, the classification are also evaluated in different time windows using different extracted time series. Finally, an optimal fusion method was also applied to estimate the outcomes using a combination of time series.

6.2.5.1. RPS's Parameters adjustment

The time delay τ and the embedding dimension M are chosen based on methods described in section 6.2.2 (figure 6.3). Based on these measures, we found that $M=3$ and $\tau=3$ are the mean of the best values for these 2 parameters over the whole population. Figure 6.4 shows examples of RR time series and its reconstructed phase space for negative and positive subjects.

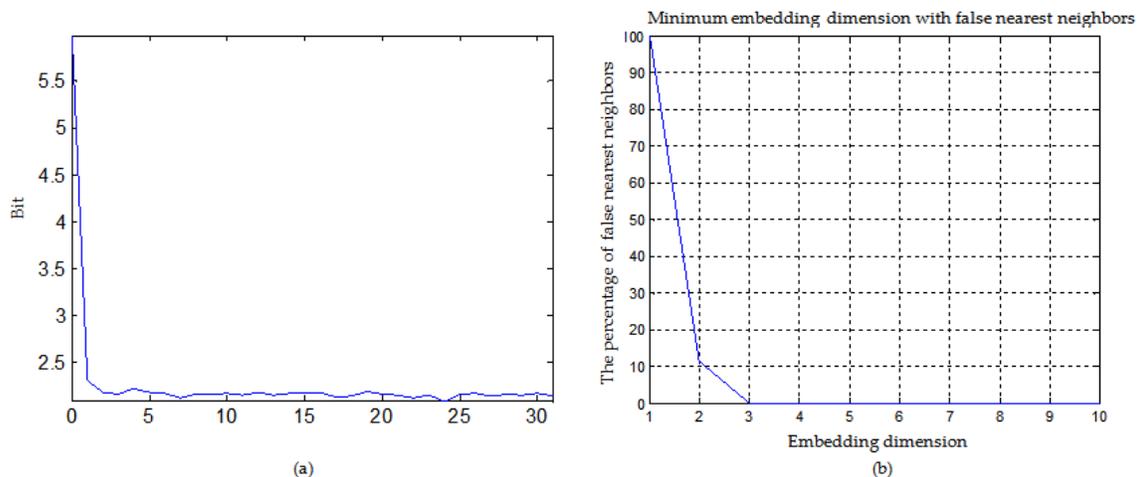


FIG. 6.3 - Mutual information representation across different time delays (a) and the percentage of false nearest neighbors across different embedding dimensions (b).

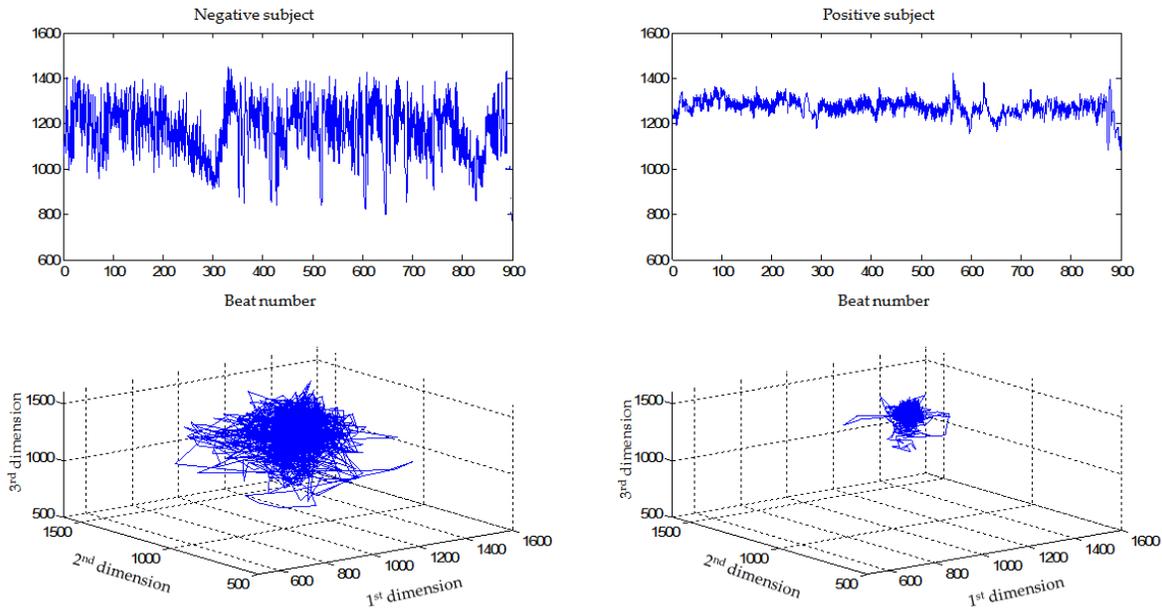


FIG. 6.4 - RR time series and its reconstructed phase space of for negative (left panel) and positive (right panel) subjects with $M=3$ and $\tau=3$.

6.2.5.2. Classification performance

Statistical analysis was performed using Statistics Toolbox of Matlab 2011a to compare the values of the computed parameters between negative and positive test groups. In this inter-group comparison evaluated by Mann-Whitney test, a p -value <0.05 was considered as significant. As regard to the density distribution in cells, the parameters that displayed a statistical difference are: the density of points in the cells number 588, 589, 859, 969 (with p -value respectively: 0.004; 0.0065; 0.0013 and 0.0036), details about these cells are reported in table 6.3. Examples of histograms based on the distribution of points in each cell are presented in figure 6.5 and clearly show the difference during the rest position between negative and positive subjects.

TAB. 6.3 - The bounds of the significant cells and their neighbors.

Cell number	Bounds in 1st axis	Bounds in 2nd axis	Bounds in 3rd axis
587	0.5-0.6	0.8-0.9	0.6-0.7
588	0.5-0.6	0.8-0.9	0.7-0.8
589	0.5-0.6	0.8-0.9	0.8-0.9
858	0.8-0.9	0.5-0.6	0.7-0.8
859	0.8-0.9	0.5-0.6	0.8-0.9
969	0.9-1	0.6-0.7	0.8-0.9
967	0.9-1	0.6-0.7	0.9-1

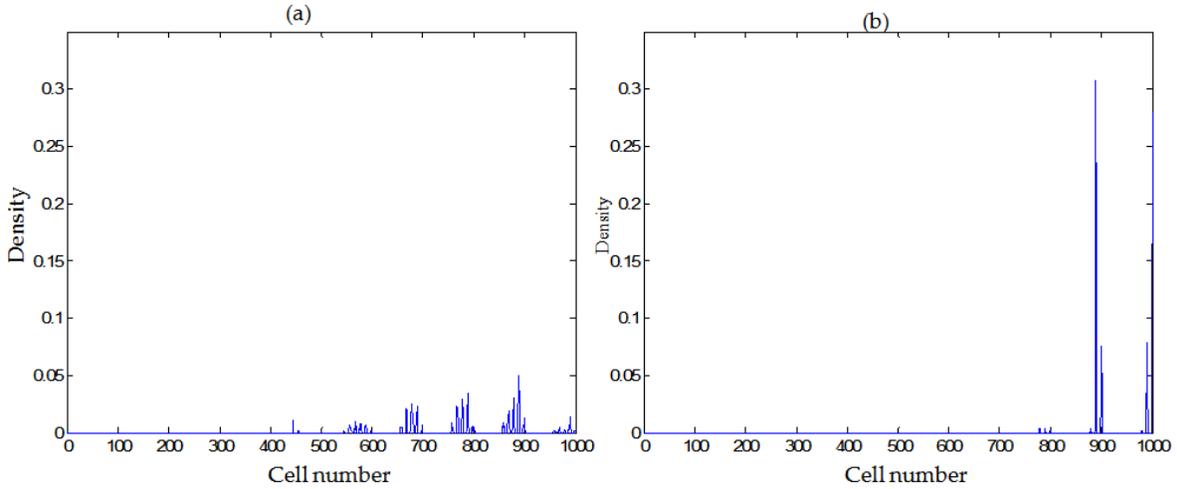


FIG.6.5 - Phase space density plot of (a) negative subject (b) positive subject.

Moreover none of the RQA parameters have shown significant differences between the two groups, but we observed that 2 parameters reflect some trends ($R_e R_a$ and TT) with p -value which are near the significance threshold (table 6.4). Parameters showing statistical difference (p -value <0.05) between the two groups were used to participate in the non-linear classification tests using kernel support vector machine (KSVM) with Gaussian radial basis kernel as kernel function.

TAB. 6.4 - Mean and standard deviation of RQA parameters.

	Negative	Positive	p-value
$R_e R_a$	0,0018 \pm 0,0014	0,0022 \pm 0,0013	0,061 ⁺
DET	0,034 \pm 0,031	0,039 \pm 0,041	0,657
$\langle L \rangle$	2,013 \pm 0,397	2,044 \pm 0,386	0,479
L_{max}	3,322 \pm 2,344	4,171 \pm 4,409	0,513
$ENTR$	0,146 \pm 0,209	0,193 \pm 0,227	0,354
LAM	0,027 \pm 0,025	0,049 \pm 0,061	0,136
TT	2,076 \pm 0,261	2,093 \pm 0,187	0,099 ⁺
V_{max}	2,452 \pm 1,207	2,886 \pm 1,676	0,126
RTE	0,789 \pm 0,043	0,782 \pm 0,0371	0,126

An exhaustive search for all possible combinations of the significant parameters was done to find an optimal performance with minimum error of classification of subjects. We remind here that we would like to propose a classical test where all positive subjects are correctly classified. Kernel parameter (σ) was set based on the value that gives the best performance with LOOCV. By applying a cross validation procedure repeated 10 times using 1/3 of the population in the training step, we determined the capability of correctly classifying the majority of positive patients. A sensitivity of 95%

with specificity of 47% was observed using the density of points in the 4 cells showing statistical difference between the two groups when the σ is equal 0.5. Table 6.5 shows the confusion matrix obtained from the average result of the 10 iterations on the test set.

TAB 6.5 - Matrix representing the average result of the 10 iterations.

	Classified as	
	Negative	Positive
Negative	9.4	10.6
Positive	1.3	23.7

6.2.5.3. Performance using PLS

Statistically significant parameters were also analyzed using partial least square regression (see the appendix of chapter 5 for further details) where 50% of the subjects are used for the training step and the rest for the testing step. Figure 6.6 and 6.7 show respectively the quality of the model which is very high on the x components and the correlation with t1 and t2 in learning phase and figure 6.8 illustrates the ROC curve of the testing data, which verifies that at around 50% of specificity we can find a high value of sensitivity. We would like to underline that the ROC curve is clearly above the random decision (black line in figure 6.8).

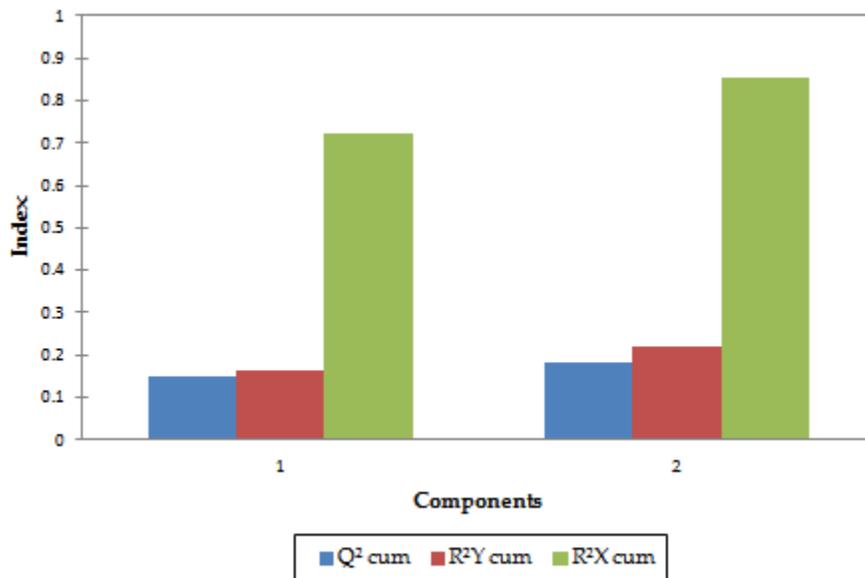


FIG. 6.6 - representation of the model quality by number of components

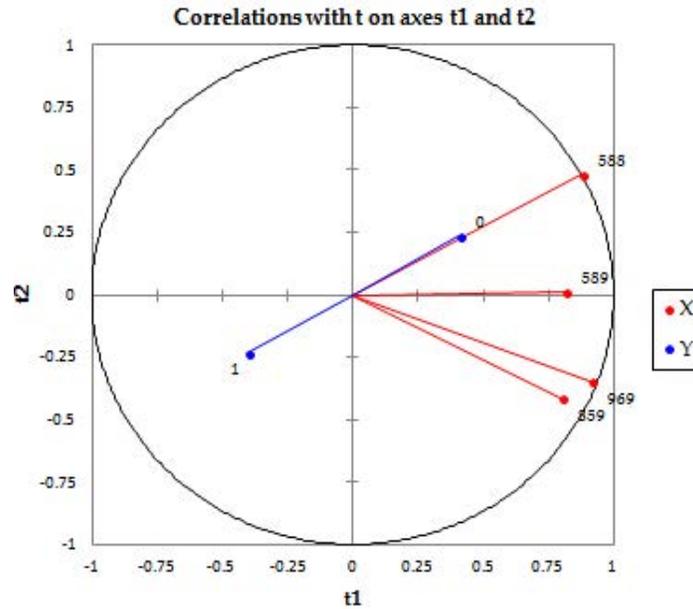


FIG. 6.7 - correlation with t1 and t2

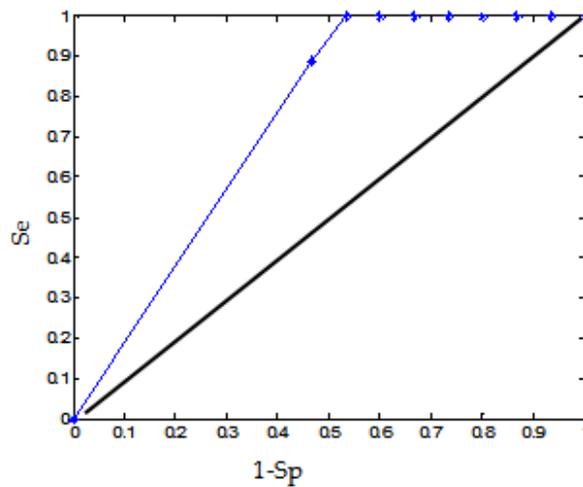


FIG. 6.8 - ROC curve of the testing phase

6.2.5.4. Performance of other time series in different time-windows

The phase space density and recurrence quantification analysis parameters were also computed from the time series described in chapter 5 (*PTT*, *AmpS* and *dPdt_max*) in the different time windows (Supine5, Early5, Last5 and Early15). Here, only the 57 subjects who have proper blood pressure signals are included. Table 6.6 summarizes the results obtained using different combinations of features selection and classification methods (Relief-KSVM; Relief-KNN; SFS-KSVM; SFS-KNN; *Probe* feature-KSVM; *Probe* feature-KNN) while minimizing the probability error of classification, it reports the best

performance measured among all combinations in each condition for each time series, it also reports the value of τ and M used in each case. Performance resulting from time series extracted from blood pressure is obtained using 1/3 of subjects for the training and 2/3 for the testing step. Table 6.6 shows that the best performance in Supine5 is with Se=69.6% and Sp=68.9% using *AmpS*; In Early5 all the time series give approximately a sensitivity and specificity around 55%, whereas in Last5 the performance is enhanced and reaches Se=89% and Sp=70% using *dPdt_max* and Se=50% and Sp=91.03% with *AmpS*. Although the values of Se and Sp are generally high in Last5 but they are not feasible in term of early prediction (the prediction is in the last 5 min before syncope), when the time-window is enlarged after tilting position (Early15) we can find performance near to that obtained in Last5 (Se=82.3%; Sp=67.7%) using RR time series. Applying the combination of SFS-KSVM on the RR time series during the 12 min of supine position (Supine12) provides the best performance and the selected features are the same which gave statistical difference (section 6.2.5.2) between the two groups (in exception the cell number 969). However the other feature selection algorithms did not select these features and the results were not better than those using SFS-KSVM combination. It is worth mention that the majority of subjects present improper blood pressure during the supine period (due to recalibration of the blood pressure) which prevented us to analyze time series extracted from this signal during the whole supine period.

TAB. 6.6 - Optimal classification results obtained using parameters from different time series, searching for minimum error of classification.

Time series	Condition	M	τ	Se	Sp	Parameters
<i>PTT</i>	Supine5	3	3	67.5%	59.7%	<i>DET; LAM; ENTR; RPDE</i>
	Early5	3	4	53.2%	57.6%	<i>DET; cells number: 899; 990</i>
	Last5	3	3	61.1%	70.7%	<i>LAM; cell number: 889</i>
	Early15	3	6	75%	73.4%	<i>V_{max}; DET; cell number: 900</i>
<i>AmpS</i>	Supine5	3	4	69.6%	68.9%	<i>DET; cells number: 38; 41</i>
	Early5	3	4	63.2%	46.5%	Cells number: 578; 659; 758; 557; 756
	Last5	3	4	50%	91.03%	<i>LAM; cell number: 678</i>
	Early15	3	6	61.4%	73.1%	<i><L>; RPDE; cell number: 677; 888</i>
<i>dPdt_max</i>	Supine5	3	4	63,6%	63.8%	Cells number: 877; 989
	Early5	3	4	56.4%	56.9%	<i>DET; cell number: 677; 890; 989</i>
	Last5	3	4	89.6%	70.3%	<i>V_{max}; cell number: 680; 987</i>
	Early15	3	6	54.3%	71.7%	<i>DET; cells number: 879; 888; 968; 998</i>
RR	Supine5	3	3	50%	51.6%	<i>DET; cell number: 978</i>
	Early5	3	4	56.6%	56.8%	Cell number: 780
	Last5	3	4	73.7%	82.9%	<i><L>; cell number: 680</i>
	Early15	3	7	82.3%	67.7%	<i>ReRa; cell number: 680; 800; 870</i>
	Supine 12	3	3	85.1%	60.9%	Cells number: 810; 588; 589; 859

6.2.5.5. Optimal fusion decision rule

The last paragraph has shown that the different time series extracted from pressure signals are relevant for our classification purpose. To continue this issue, an optimal data fusion technique (Hernandez et al., 1999) was applied to evaluate the performance of the different pressure time series combination. The objective is to test the following two hypotheses:

- H_0 : no syncope will occur during the test, the subject belongs to the negative group.
- H_1 : the patient will be subject to syncope, so he belongs to the positive group.

Thus, a Boolean output can be obtained from local decisions u_i such as: under the hypothesis H_0 : the decision $u_i = -1$ and under the hypothesis H_1 : $u_i = 1$. The fusion problem produces a new statistical decision $u = f(u_1, \dots, u_n)$ based on the association of a set of local decisions. The function f determined by Chair and Varshney (Z. Chair, 1986) is a weighted sum of local decisions (derived from different sources, here they are derived from different time series) in which the weights are defined as a function of the detection performance of each individual time series:

$$f(u_1, \dots, u_n) = \begin{cases} 1, & a_0 + \sum_{i=1}^n a_i u_i > 0 \\ -1, & \text{otherwise} \end{cases} \quad (6.12)$$

Where the optimum weights a_i are given by:

$$a_0 = \log \frac{P_1}{P_0} ,$$

$$a_i = \log \frac{1 - P_{Mi}}{P_{Fi}} , \text{ if } u_i = +1$$

$$a_i = \log \frac{1 - P_{Fi}}{P_{Mi}} , \text{ if } u_i = -1$$

$P_1 = P(H_1)$ and $P_0 = P(H_0)$ are a priori probabilities of the two hypotheses; P_{Fi} and P_{Mi} are respectively the probability of false alarm and miss for the i^{th} time series. Figure 6.9 shows the structure of this optimal fusion rule used in this chapter. First we combined time series extracted from pressure signal only, and then we add the RR-interval.

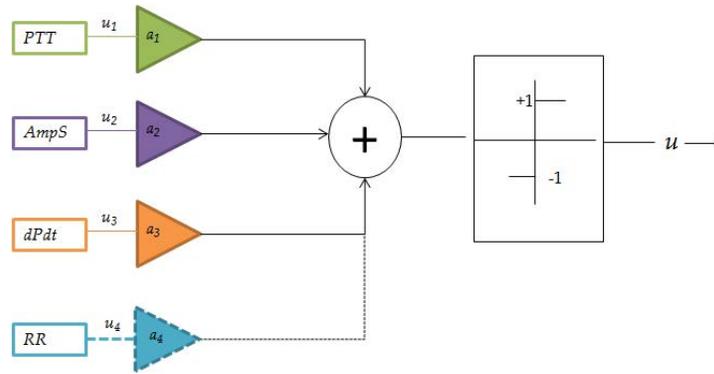


FIG. 6.9 - Optimal fusion rule applied to time series extracted from pressure signal.

Table 6.7 represents the results of the optimal fusion in different time windows. We retained here from each sensor a $P_{Mi} = (1-Sp)$ and $P_{Fi} = (1-Se)$ of table 6.6 which were established for a minimum probability of error. These results show that the performance was improved in term of sensitivity and specificity in the different time windows. The intervention of RR-interval time series help to improve only the specificity in Early5.

TAB. 6.7- Optimal fusion results in different time windows using first only the time series extracted from pressure signal then the time series from ECG and pressure signals

Optimal fusion	Time series from pressure signal		Time series from ECG and pressure signals	
	Se	Sp	Se	Sp
Supine5	100%	86.7%	100%	78.6%
Early5	64.3%	66.7%	64.3%	80%
Last5	85.7%	93.3%	85.7%	93.3%
Early15	92.8%	100%	92.8%	93.3%

6.3. Hidden Semi-Markov models

Hidden Semi-Markov models (Yamagishi et al., 2006) are a specialization of hidden Markov models that provide an explicit representation of state durations. In the classic HMM, the possibility of a self-transition to the current state (which is not present on HSMMs), implies a geometrical probability density function (PDF) for the duration on each state. The geometrical PDF is strictly decreasing in time and is thus not adapted to the representation of the biomedical time series of interest in our work. In contrast, the PDF of state durations in an HSMM can be explicitly defined, as a function of the problem being addressed; this difference between a state in HMM and in HSMM is exposed in figure 6.10 (more details about HMM and HSMM are reported in appendix C). In this chapter, we have chosen an explicit-duration HSMM with a Gaussian law, truncated in zero, to represent state durations.

Concerning the representation of the observations, since the time series analyzed in this work represent continuous variables, a continuous HSMM will be used. Observation distributions of each state will be thus defined by Gaussians, each one representing a subspace of the dynamic range of the available observations, as described in (Rabiner, 1989), (Yu, 2010). By using the above-mentioned structure, each HSMM (M_k) is defined by one hyper-parameter representing the number of states of the model (S_k) and the set of model parameters:

$$\theta_k = \{\pi_i, a_{ij}, b_i(\vec{\mu}_i, \Sigma_i), p_i(\mu_{i,d}, \sigma_{i,d})\} \quad (6.13)$$

Where, for each state i , π_i are the initial probabilities, a_{ij} are the transition matrix coefficients (with $a_{i,i}=0$), $b_i(\vec{\mu}_i, \Sigma_i)$ are the multivariate Gaussian laws representing the observation space ($\vec{\mu}_i$ is the mean vector and Σ_i is the covariance matrix) and $p_i(\mu_{i,d}, \sigma_{i,d})$ are the Gaussian laws representing the duration on each state ($\mu_{i,d}$ and $\sigma_{i,d}$ are respectively the mean and the standard deviation).

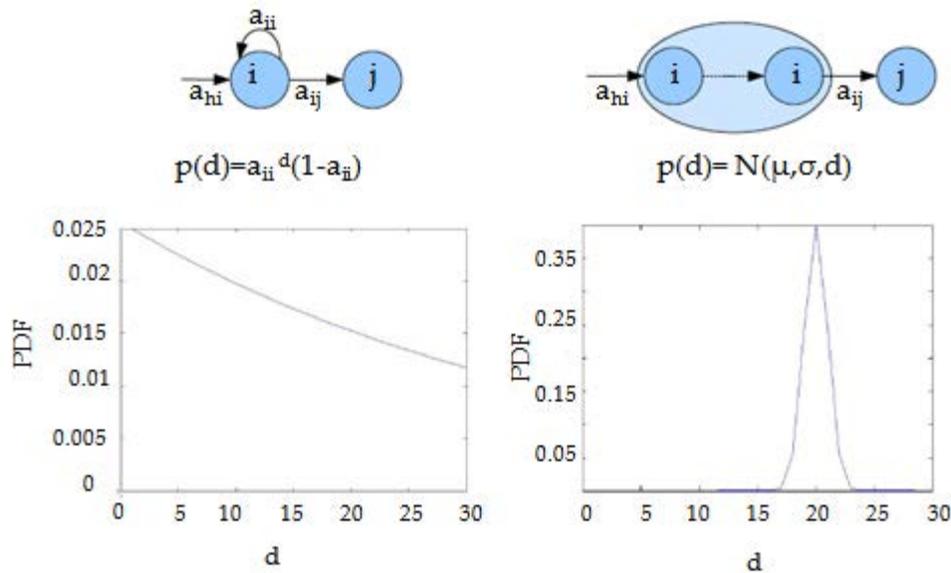


FIG.6.10- Difference between HMM / HSMM on the modeling of the time spent in a state: in the left panel a state in HMM whose PDF of duration follows a geometrical distribution function, in the right panel a state of HSMM of which the PDF of duration follows a normal distribution with μ and σ parameters.

6.3.1. Parameter estimation of the HSMM

Parameter S_k is firstly estimated with the Bayesian information criterion (BIC)(Schwarz, 1978). Then, a learning phase is applied to each HSMM in order to estimate θ_k . This phase consists in finding the optimal set of parameters θ_k^* that maximizes the likelihood criteria:

$$\theta_k^* = \arg \max_{\theta} \prod_{r=1}^R P(O^r / M_k(\theta_k)) \quad (6.14)$$

Where $P(O^r / M_k(\theta_k))$ is the probability that the observation sequence O^r , $r \in \{1, \dots, R\}$, is generated by model M_k with parameters θ_k and R is the number of observed time series available for the learning phase. This optimization is carried out by a two-stage process:

- *Initialization*: The initial values of $\vec{\mu}_i$ and Σ_i are estimated by applying a spatial clustering method based on Gaussian Mixtures Models (GMM) including S_k components. Each component of the GMM will thus correspond to the initial observation distribution for each state i of the HSMM.
- *Learning phase*: In this phase, the model parameters are estimated through an Expectation-Maximization (EM) algorithm, focused on the most probable state path, obtained with the Viterbi algorithm (Forney, 1973). The applied EM loop can be decomposed into three steps:
 - Step E1 (Estimation, forward phase): the probability of being in state i at any time t ($L_t(i)$), is estimated with the following equation

$$L_t(i) = \max_{1 \leq d \leq D} \left(p_i(d) \prod_{t-d+1}^t b_j(O_t^r) \max_{1 \leq j \leq S_k} (L_{t-d}(j) a_{ji}) \right) \quad (6.15)$$

Where $t \in \{D, D+1, \dots, T-1, T\}$, d is the duration on state i and D is the maximal duration supported in the optimization phase. $L_t(i)$ is thus maximized with respect to the previous state j and its duration d . For each i and t , the values of d and j maximizing $L_t(i)$ ($d_{i,t}^*$ and $j_{i,t}^*$ respectively) are recorded in order to be used during the backward phase.

- Step E2 (Estimation, backward phase): starting from $q_N = j_{i,T}^*$, a backward search algorithm is applied until $t=0$ to recursively estimate the sequence of states $Q = \{q_1, q_2, \dots, q_N\}$ that maximizes $P(O^r / M_k(\theta_k))$. A number of intermediate variables are computed during this process, to be used in the maximization phase. Steps E1 and E2 are applied to the R time series available on the learning set.
- Step M (Maximization): parameters π_i , a_{ij} , $\vec{\mu}_i$, Σ_i , $\mu_{i,d}$ and $\sigma_{i,d}$ are updated according to the variables estimated during phases E1 and E2.

The EM algorithm is applied until a convergence criterion is achieved or until a maximum number of iterations is reached. More details on this algorithm can be found in (Yu, 2010) and the references therein.

6.3.2. Syncope prediction process

The proposed HSMM-based prediction of syncope has been developed in three phases:

1. ECG and blood pressure signal processing is done in order to extract a set of features from each detected beat and pulse and to build time series from each extracted feature,
2. Training of two HSMM (Mod_{neg} by the negative subjects and Mod_{pos} by the positive subjects) using data from a learning dataset (DB_L),
3. Evaluation of the detection performance on a test dataset (DB_T), based on the classification of an observation from the maximum likelihood of models Mod_{neg} and Mod_{pos} .

These steps are described by the next subsections.

6.3.2.1. Feature extraction and preparation

The time series of the RR-interval extracted from the ECG signal, the amplitude of the systolic blood pressure (Amp_S), the amplitude of the first derivative systolic blood pressure ($dPdt_{max}$) and the variability of the pulse transit time (PTT) previously described are used here. Each time series is resampled at 2 Hz and pre-processed by eliminating the average value of the first 100 beats. This approach allows for a differentiation of time series in priority according to their dynamics and not according to their absolute amplitudes. After that, a temporal window that includes the most important transition event that happened when the table is tilted is extracted to be used to learn or test the models (figure 6.11).

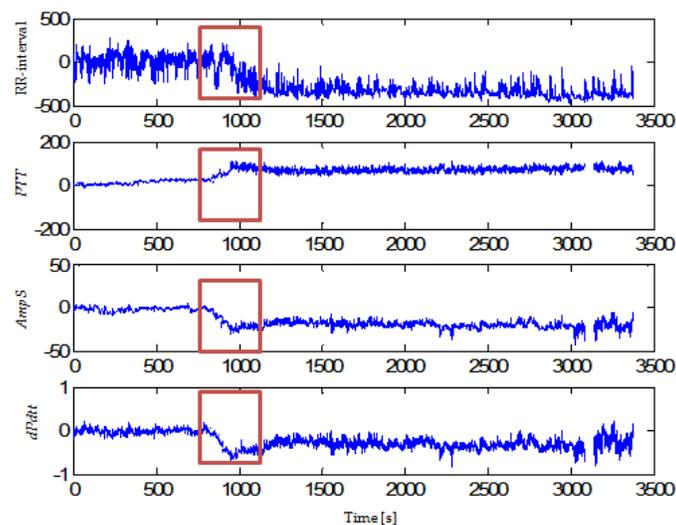


FIG. 6.11- Time series of the extracted features and illustration of the window in red frame that contains the transition event that happened when the table is tilted.

6.3.2.2. Training and testing the two HSMM models

As already mentioned, two different HSMM are learned in order to differentiate the dynamics of the negative and the positive test group. The learning phase is firstly applied to Mod_{neg} . The number of states of this model $S_{k,neg}$ is estimated and then the set of model parameters $\theta_{k,neg}$ are obtained, as described in section 6.3.1. The positive model Mod_{pos} inherits the number of states from Mod_{neg} , $S_{k,pos} = S_{k,neg}$ and specific parameters are estimated. The window containing the transition event of each subjects of the learning database is used to learn the model.

In the test phase, the log-likelihood of each model is calculated for the transition window of each subject of the test dataset. A subject is assigned to the model for which the maximum likelihood is obtained.

6.3.3. Evaluation of the algorithm

The two learning datasets ($DB_{L,neg}$ and $DB_{L,pos}$) representing respectively negative and positive groups are extracted from the database already described in previous chapter, to create specific HSMM using a random selection of half of the subjects in database. The algorithm is evaluated on test datasets ($DB_{T,neg}$ and $DB_{T,pos}$) composed of the remaining subjects available in the database, but not included in the learning datasets. Ten different realizations of these learning and test databases were generated.

As mentioned above, only the window containing the transition event (when the table is tilted) is used. After training the two models, Mod_{neg} and Mod_{pos} by $DB_{L,neg}$ and $DB_{L,pos}$, the performance is evaluated on the test dataset $DB_{T,neg}$ and $DB_{T,pos}$. Table 6.8 summarizes the optimal results obtained using time series of different features and table 6.9 reports the performance when we use combination of the time series of all features together as described in section 6.2.5.5.

TAB - 6.8. Optimal results obtained by block detection approach

Feature	Nb_states (S_k)	Accuracy	Sensitivity	Specificity
RR-interval	30	57.2%	67.8%	46.4%
PTT	30	51%	56%	46%
AmpS	25	49.3%	53.6%	45%
dPdt_max	35	47.5%	51.4%	43.6%

TAB - 6.9. Results of optimal fusion using first only the time series extracted from pressure signal then the time series from ECG and pressure signals.

	Sensitivity	Specificity
Fusion without RR-interval time series	78.6%	50%
Fusion with RR-interval time series	64.3%	64.3%

Again, we observed the data fusion approach improves the classification of each local detector but it is worth to mention that the performances do not reach the accuracy of the reconstructed phase space.

6.4. Conclusion

In the present chapter, two non-linear dynamic methods were evaluated. First, a novel method is proposed to evaluate the dynamic behavior of time series and to distinguish between patients with negative and positive response to head up tilt test, using parameters extracted from the reconstructed phase space of these time series. Applying these proposed parameters, we achieved a correct classification on most of the positive patients during the supine position. Although the obtained specificity is not very high, our findings are interesting in terms of sensitivity leading to predict the occurrence of syncope before tilted position and in terms of the identification of features containing useful information for HUTT processing.

The phase space reconstruction combined with classification process demonstrates the interest to take into account the dynamics of the RR series and their capability to predict tilt test's outcome using only pre-HUTT data. We also performed RPS analysis on the time series extracted from pressure signals (*AmpS*, *dPdt_max*, *PTT*) and combined their results through an optimal fusion node. The optimal classification results obtained underlines the importance of taking into consideration not only the dynamics of RR time series but also the dynamics of other cardiovascular time series. To our knowledge, and compared to the results reported in the literature, the new tool used in this application provides promising results, and can be seen complementary to those obtained using nonlinear parameters (chapter 4) and coherence in time frequency domain (chapter 5) during the 15 min following the tilt position.

Second, an evaluation of the classification of the two groups of subjects based on HSMM is proposed. HSMM are applied to learn the difference of dynamics between negative and positive test groups of various time series extracted from the ECG and the blood pressure signals. An EM algorithm, itself based on the Viterbi algorithm has been extended to the HSMM to learn its various parameters. With this approach the dynamics could be efficiently represented by the parameters of the models, with limited computation time and reduced number of hyper-parameters to tune. To test new time series, the two models are put in competition and comparison between the likelihoods obtained is done to provide a decision. Time series of several features are used with several approaches to take the decision. Unfortunately, the results show weak performance and all approaches did not provide satisfactory results. This could be related to the heterogeneous dynamics presented in each test group. However, ultimate improvements can be done by changing the size of analyzed windows, using more data to learn the models or applying specific transformation of the data series before use or even using more complex HSMM architectures, such as coupled hidden Markov models.

Appendix C

Markov and Hidden Semi-Markov models

In this section, a brief introduction to Hidden Markov Model (HMM) and Hidden Semi-Markov Model (HSMM) and their based properties is presented. For more details, we invite the reader to refer to the following references, (Ephraim and Merhav, 2002) or (MacDonald and Zucchini, 1997) for HMM and (Kulkarni, 1996) or (Yu, 2010) for HSMM.

An HMM is a double stochastic process in which the underlying stochastic process is a Markov chain with finite state and discrete time. The sequence of states is not observable (hence the term "hidden") and influences another stochastic process that produces a sequence of observations. The basic structure of the HMM is illustrated in figure C.1.

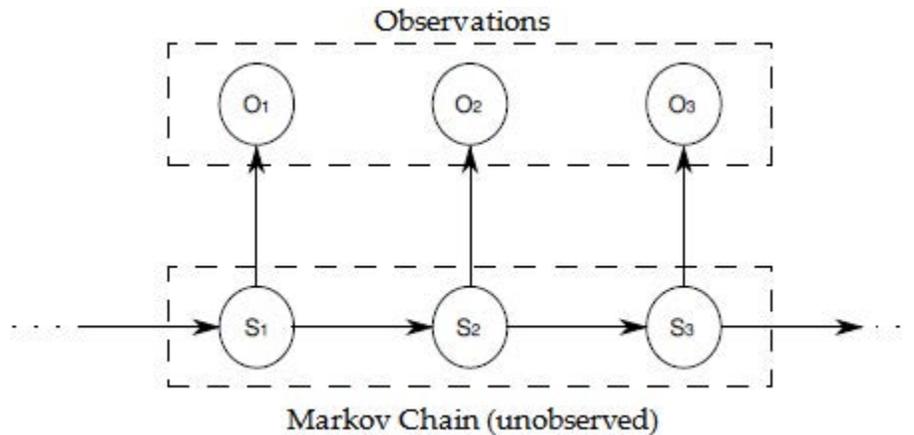


FIG. C.1 - Basic structure of a hidden Markov model (Altuve, 2011).

MMC was applied in several domains such as biology (Munshaw and Kepler, 2010), speech recognition (Rabiner, 1989), image processing (Yamato et al., 1992) and text recognition (Chen et al., 1994). However, due to a non-zero probability of self-transition of a state, the duration of a state in HMM follows a geometric distribution. This makes the HMM inadequate in some applications, such as the recognition of human genes in DNA or financial modeling.

To solve this problem, FERGUSON (Ferguson, 1980) proposed a model that allows arbitrary time distributions of staying for hidden process. As the hidden process

becomes semi-Markov, this model is called a hidden semi-Markov⁶ model. The basic structure of HSMM is illustrated in figure C.2. Different hypotheses have been considered in the literature for the distribution of observations, as well as the representation of time spent in a state. For the latter three types of modeling are mainly used: a parametric family of continuous distributions (Levinson, 1986), a family of exponential distributions (Mitchell and Jamieson, 1993) or the negative binomial law (Richard Durbin, 1998).

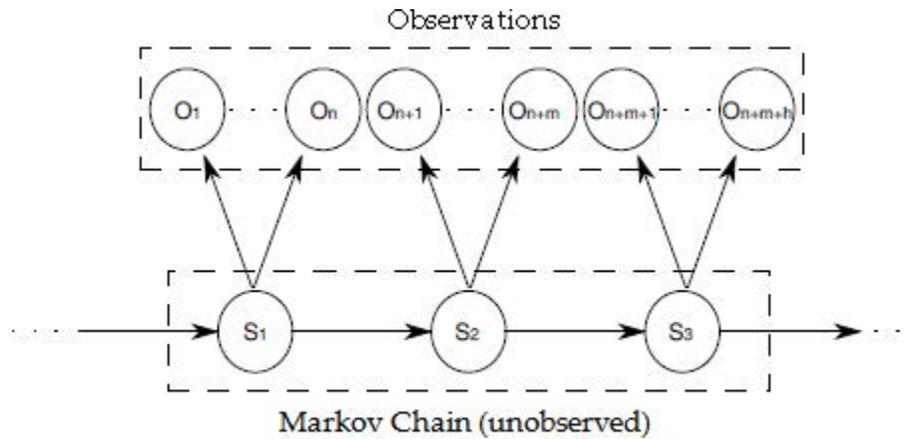


FIG. C.2 - Basic structure of a hidden Markov model.

Models architecture

Let $S = \{1, \dots, M\}$ the set of states of a Markov chain. The sequence of states is designated by $S_{1:T} = S_1, \dots, S_T$, where $S_t \in S$ is the state at time t . A realization of $S_{1:T}$ is denoted by $s_{1:T}$.

The observation sequence is often denoted by $O_{1:T} = O_1, \dots, O_T$, where $O_t \in \mathcal{V}$ is the observation at time t and $\mathcal{V} = \{v_1, v_2, \dots, v_k\}$ is the set of observable values.

In HMM, an observation is generated in each state i because the states do not have duration. For example, for the sequence of observation $o_{1:T}$, the sequence of

underlying Markov states is $s_{1:T} = i_1, i_1, i_1, i_2, \dots, i_n$ where $\sum_{m=1}^n s_m = T$ and $s_m = T$ and $i_1, i_2, \dots, i_n \in S$.

In contrast, in HSMM, each state i has a variable duration, which is associated to the number of observations produced when remains in the state. The duration of the

⁶ A semi-Markov process has the following property: the probability of moving from a hidden state i to another j depends only on the time spent since the entry into the current state i . Contrariwise, the property of a Markov process assumes that the probability of changing state depends only on the current state.

states is a random variable and takes an integer value in the set $D = \{1, 2, \dots, D\}$. For example, for the sequence of observation $o_{1:T}$, the sequence of semi-Markovian states underlying is $s_{1:T} = i_1, i_2, \dots, i_n$ and the sequence of duration of states is $d_{1:T} = d_1, d_2, \dots, d_n$,

where $\sum_{m=1}^n d_m = T$, $i_1, i_2, \dots, i_n \in S$ and $d_1, d_2, \dots, d_n \in D$.

The probability of transition between states $i \rightarrow j, \forall i \neq j$, is defined by:

$$a_{ij} = P(S_{t+1} = j | S_t = i), \quad (C.1)$$

With $\sum_{j \in S \setminus \{i\}} a_{ij} = 1$ and the auto-transition probability $a_{ii} \neq 0$ for HMM and $a_{ii} = 0$ for HSMM, where $i, j \in S$.

In HSMM, the duration of a state is independent from the duration of the previous state and is conditioned only on the current state. Therefore, the probability of duration of the state in the HSMM is denoted by:

$$p_i(d) = P(S_{t+d-1} = S_t = i) \quad (C.2)$$

As mentioned before, HMM does not have a defined probability of duration.

The probability of emission observations of the state i is defined by:

$$b_i(v_k) = P(o_t = v_k | S_t = i), \quad (C.3)$$

With $\sum_{v_k} b_i(v_k) = 1$ and the initial probability of a state is defined by:

$$\pi_i = P(S_1 = i), \quad (C.4)$$

With $\sum_i \pi_i = 1$.

Thus, the set of parameters λ^{HMM} of HMM is defined by:

$$\lambda^{HMM} = \{a_{ij}, b_i(v_k), \pi_i\}. \quad (C.5)$$

And the set of parameters λ^{HSMM} of HSMM is defined by:

$$\lambda^{HSMM} = \{a_{ij}, b_i(v_k), \pi_i, p_i(d)\}. \quad (C.6)$$

Table C.1 summarizes the main parameters and characteristics of the HMM and HSMM and summarizes the notation used in this manuscript.

An example of an HMM is shown in Figure C.3. The first state i_1 is selected according to the probability π_1 . In the state i_1 , the observation o_1 is produced according to the probability of emission of observation $b_{i_1}(o_1)$. It transits (stays) to the state i_1 following the probability of transition a_{11} and produces the observation o_2 according to

probability $b_{i_1}(o_2)$. Then it passes to i_2 based on a_{12} and produces o_3 according to $b_{i_2}(o_3)$. Then it transits to $i_2 \dots i_n$ until the last observation occurred.

TAB. C.1 - Summary of the main parameters and characteristics of the HMM and HSMM.

Parameter	Symbol	HMM	HSMM
element of the transition matrix	a_{ij}	$\sum_{j \in S \setminus \{i\}} a_{ij} = 1$	$a_{ii} \neq 0$
Probability of observation	$b_i(v_k)$	$\sum_{v_k} b_i(v_k) = 1$	$\sum_{v_k} b_i(v_k) = 1$
Probability of duration	$p_i(d)$	Not defined	Gaussian law
Initial probability	π_i	$\sum_i \pi_i = 1$	$\sum_i \pi_i = 1$

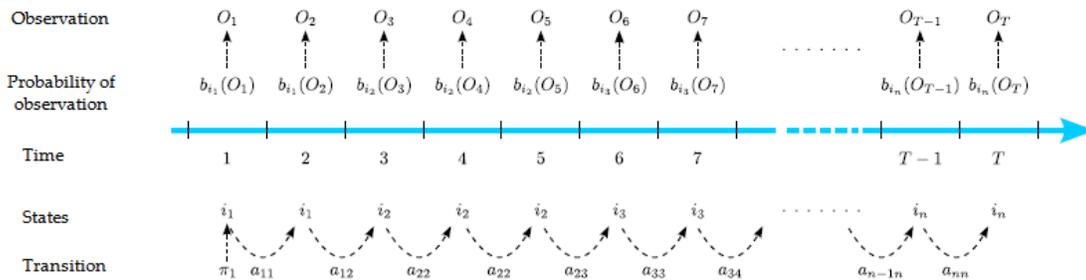


FIG. C.3 – Principle of hidden Markov Model (Altuve, 2011).

An example of an HSMM, inspired by YU [Yu, 2010], is shown in Figure C.4. The first state i_1 is chosen according to the probability π_1 . The process continues for a period $d_1 = 2$ in the state and i_1 and two observations are thus produced in this time interval $\langle o_1, o_2 \rangle$, following the probabilities of emission of observations $\langle b_{i_1}(o_1); b_{i_1}(o_2) \rangle$. It transits to the state i_2 based on the probability of transition a_{12} where it will stay $d = 3$ time units and produce three observations $\langle o_3, o_4, o_5 \rangle$ according to the probabilities of emission of observations $\langle b_{i_2}(o_3), b_{i_2}(o_4), b_{i_2}(o_5) \rangle$. Then it transits to i_3, \dots, i_n until the last observation occurred.

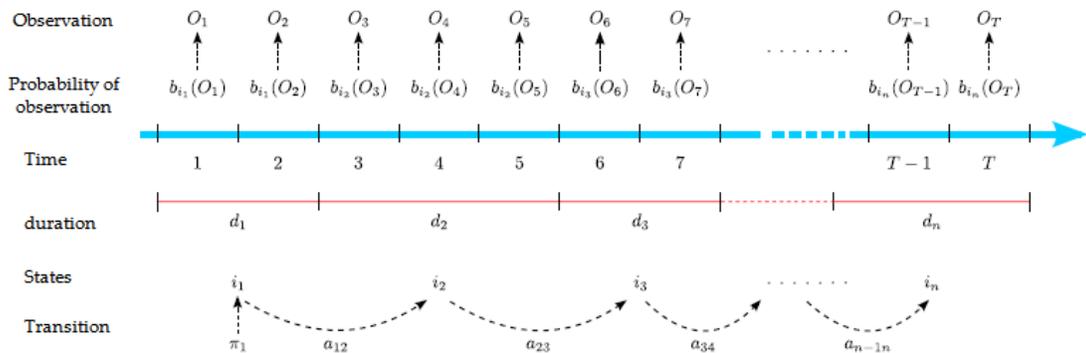


FIG. C.4 – Principle of hidden semi-Markov Model (Altuve, 2011).

References

- Abarbanel, H., 1997. *Analysis of Observed Chaotic Data*. Springer New York.
- Altuve M., 2011. Détection multivariée des épisodes d'apnée-bradycardie chez le prématuré par modèles semi-Markovien cachés.
- Argyris, J.H., Faust, G., Haase, M., 1994. *An Exploration of Chaos: An Introduction for Natural Scientists and Engineers*. North Holland, Amsterdam ; New York.
- Buzug, T., Pfister, G., 1992. Comparison of algorithms calculating optimal embedding parameters for delay time coordinates. *Phys. Nonlinear Phenom.* 58, 127–137. doi:10.1016/0167-2789(92)90104-U
- Cassidy, M.J., Penny, W.D., 2002. Bayesian nonstationary autoregressive models for biomedical signal analysis. *IEEE Trans. Biomed. Eng.* 49, 1142–1152.
- Cellucci, C.J., Albano, A.M., Rapp, P.E., 2003. Comparative study of embedding methods. *Phys. Rev. E* 67, 066210. doi:10.1103/PhysRevE.67.066210
- Chen, M.-Y., Kundu, A., Zhou, J., 1994. Off-line handwritten word recognition using a hidden Markov model type stochastic network. *IEEE Trans. Pattern Anal. Mach. Intell.* 16, 481–496. doi:10.1109/34.291449
- Eckmann, J.-P., Kamphorst, S.O., Ruelle, D., 1987. Recurrence Plots of Dynamical Systems. *EPL Europhys. Lett.* 4, 973. doi:10.1209/0295-5075/4/9/004
- Ephraim, Y., Merhav, N., 2002. Hidden Markov processes. *IEEE Trans. Inf. Theory* 48, 1518–1569. doi:10.1109/TIT.2002.1003838
- Ferguson, Jack D. Variable duration models for speech. In *Proc. Symposium on the application of hidden Markov models to text and speech*, pp. 143-179. 1980
- Faisan, S., Thoraval, L., Armspach, J.-P., Metz-Lutz, M.-N., Heitz, F., 2005. Unsupervised learning and mapping of active brain functional MRI signals based on hidden semi-Markov event sequence models. *IEEE Trans. Med. Imaging* 24, 263–276.
- Forney, J., G.D., 1973. The viterbi algorithm. *Proc. IEEE* 61, 268–278. doi:10.1109/PROC.1973.9030
- Fraser, A.M., Swinney, H.L., 1986. Independent coordinates for strange attractors from mutual information. *Phys. Rev. A* 33, 1134–1140. doi:10.1103/PhysRevA.33.1134
- Gao, J., Cai, H., 2000. On the Structures and Quantification of Recurrence Plots.
- Gordon, K., Smith, A.F.M., 1990. Modeling and Monitoring Biomedical Times Series. *J. Am. Stat. Assoc.* 85, 328–337. doi:10.2307/2289768
- Hernandez, A.I., Carrault, G., Mora, F., Thoraval, L., Passariello, G., Schleich, J.-M., 1999. Multisensor fusion for atrial and ventricular activity detection in coronary care monitoring. *IEEE Trans. Biomed. Eng.* 46, 1186–1190. doi:10.1109/10.790494
- Jerome Connor, L.E.A., 1991. Recurrent Networks and NARMA Modeling. 301–308.

- Kennel, M.B., Brown, R., Abarbanel, H.D.I., 1992. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev. A* 45, 3403–3411. doi:10.1103/PhysRevA.45.3403
- Kraskov, A., Stögbauer, H., Grassberger, P., 2004. Estimating mutual information. *Phys. Rev. E* 69, 066138. doi:10.1103/PhysRevE.69.066138
- Kulkarni, V.G., 1996. *Modeling and Analysis of Stochastic Systems*. CRC Press.
- Kumar, A., Mullick, S.K., 1996. Nonlinear dynamical analysis of speech. *J. Acoust. Soc. Am.* 100, 615–629. doi:10.1121/1.415886
- Leistritz, L., Galicki, M., Kochs, E., Zwick, E.B., Fitzek, C., Reichenbach, J.R., Witte, H., 2006. Application of generalized dynamic neural networks to biomedical data. *IEEE Trans. Biomed. Eng.* 53, 2289–2299. doi:10.1109/TBME.2006.881766
- Levinson, S.E., 1986. Continuously variable duration hidden Markov models for automatic speech recognition. *Comput. Speech Lang.* 1, 29–45. doi:10.1016/S0885-2308(86)80009-2
- MacDonald, I.L., Zucchini, W., 1997. *Hidden Markov and Other Models for Discrete-valued Time Series*. CRC Press.
- Marwan, N., Carmen Romano, M., Thiel, M., Kurths, J., 2007. Recurrence plots for the analysis of complex systems. *Phys. Rep.* 438, 237–329. doi:10.1016/j.physrep.2006.11.001
- Mitchell, C.D., Jamieson, L.H., 1993. Modeling duration in a hidden Markov model with the exponential family, in: , 1993 IEEE International Conference on Acoustics, Speech, and Signal Processing, 1993. ICASSP-93. Presented at the , 1993 IEEE International Conference on Acoustics, Speech, and Signal Processing, 1993. ICASSP-93, pp. 331–334 vol.2. doi:10.1109/ICASSP.1993.319304
- Monk, A.T., Compton, A.H., 1939. Recurrence phenomena in cosmic-ray intensity. *Rev. Mod. Phys.* 11, 173–179. doi:10.1103/RevModPhys.11.173
- Munshaw, S., Kepler, T.B., 2010. SoDA2: a Hidden Markov Model approach for identification of immunoglobulin rearrangements. *Bioinforma. Oxf. Engl.* 26, 867–872. doi:10.1093/bioinformatics/btq056
- Ott, E., 2002. *Chaos in Dynamical Systems*. Cambridge University Press.
- Petry, A., Barone, D.A.C., 2002. Speaker identification using nonlinear dynamical features. *Chaos Solitons Fractals* 13, 221–231. doi:10.1016/S0960-0779(00)00260-5
- Rabiner, L.R., 1989. A tutorial on hidden Markov models and selected applications in speech recognition, in: *PROCEEDINGS OF THE IEEE*. pp. 257–286.
- Richard Durbin, S.R.E., 1998. *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*. doi:10.1017/CBO9780511790492

- Roberto Barbera, G.L.R., 2011. Grid Computing Technology and the Recurrence Quantification Analysis to Predict Seizure Occurrence in Patients Affected by Drug-Resistant Epilepsy. *Data Driven E-Sci.* doi:10.1007/978-1-4419-8014-4_37
- Roberts, F.M., Povinelli, R.J., Ropella, K.M., 2001. Identification of ECG Arrhythmias Using Phase Space Reconstruction, in: *Proceedings of the 5th European Conference on Principles of Data Mining and Knowledge Discovery, PKDD '01.* Springer-Verlag, London, UK, UK, pp. 411–423.
- Schwarz, G., 1978. Estimating the Dimension of a Model. *Ann. Stat.* 6, 461–464. doi:10.1214/aos/1176344136
- Srinivasan, N., Wong, M.T., Krishnan, S.M., 2003. A new phase space analysis algorithm for cardiac arrhythmia detection, in: *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2003.* Presented at the Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2003, pp. 82–85 Vol.1. doi:10.1109/IEMBS.2003.1279515
- Takens, F., 1981. Detecting strange attractors in turbulence, in: Rand, D., Young, L.-S. (Eds.), *Dynamical Systems and Turbulence*, Warwick 1980, Lecture Notes in Mathematics. Springer Berlin Heidelberg, pp. 366–381.
- Tang, X., 2009. Hybrid Hidden Markov Model and Artificial Neural Network for Automatic Speech Recognition, in: *Pacific-Asia Conference on Circuits, Communications and Systems, 2009. PACCS '09.* Presented at the Pacific-Asia Conference on Circuits, Communications and Systems, 2009. PACCS '09, pp. 682–685. doi:10.1109/PACCS.2009.138
- Vásquez, C., Hernández, A., Mora, F., Carrault, G., Passariello, G., 2001. Atrial activity enhancement by Wiener filtering using an artificial neural network. *IEEE Trans. Biomed. Eng.* 48, 940–944. doi:10.1109/10.936371
- Waibel, A., Hanazawa, T., Hinton, G., Shikano, K., Lang, K.J., 1989. Phoneme recognition using time-delay neural networks. *IEEE Trans. Acoust. Speech Signal Process.* 37, 328–339. doi:10.1109/29.21701
- Yamagishi, J., Ogata, K., Nakano, Y., Isogai, J., Kobayashi, T., 2006. HSMM-Based Model Adaptation Algorithms for Average-Voice-Based Speech Synthesis, in: *2006 IEEE International Conference on Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings.* Presented at the 2006 IEEE International Conference on Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings, pp. I–I. doi:10.1109/ICASSP.2006.1659961
- Yamato, J., Ohya, J., Ishii, K., 1992. Recognizing human action in time-sequential images using hidden Markov model, in: , *1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1992. Proceedings CVPR '92.*

- Presented at the , 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1992. Proceedings CVPR '92, pp. 379–385. doi:10.1109/CVPR.1992.223161
- Yu, S.-Z., 2010. Hidden semi-Markov models. *Artif. Intell., Special Review Issue 174*, 215–243. doi:10.1016/j.artint.2009.11.011
- Z. Chair, P.K.V., 1986. Optimal Data Fusion in Multiple Sensor Detection Systems. *Aerosp. Electron. Syst. IEEE Trans. On* 98 – 101. doi:10.1109/TAES.1986.310699

Conclusion

Syncope is considered as common pathology- a majority of people passes out at least once in their lives. Even if it is not generally life threatening but it is a potentially serious problem, since any loss of consciousness can produce injury. For this reason, people experienced syncope may ask to undergo a head-up tilt test (HUTT) which is an examination procedure used to diagnose syncope. This test is based on the reproduction of the syncope's symptoms, it begins with a supine period (about 12-15 min), and then the table is tilted at an angle between 60° and 80° for duration up to 45 min. Nonetheless, its major inconvenience is the long duration of the examination procedure which can take up to one hour. Hence, reducing the examination time would decrease its economical impact and improve the comfort of the patient.

These reasons have led to the search for methods that reduce the test duration by predicting the outcomes of HUTT. However, few studies have evaluated the non-linear methods to analyze time series extracted from cardiovascular signals during the test, and have exploited their interrelation ship between HRV and SBP and their temporal dynamics. This is the challenge of this thesis, which seeks to predict the appearance of symptoms of the syncope and to differentiate between subjects with negative response (non-fainting people) and positive response (fainting people) to HUTT before the end of the test.

First, the analysis of heart rate variability-which provides information about the sympathetic and parasympathetic balance in controlling the cardiac rhythm- was studied. Conventional linear and non-linear parameters as well as baroreflex response were used and compared between two groups of people. The experimentation that carried out on 66 subjects including 35 who have developed syncope (positive patients) and 31 that did not develop syncope (negative subjects) during the tilt test, showed an improvement of the understanding of the physiological discomfort. By using a kernel based classification method (Kernel support vector machine) and the estimated nonlinear parameters jointly with some classical parameters (the long range scaling exponents of the detrended fluctuation analysis, the sample Entropy, the standard deviation and the total power spectrum) of the RR-interval time series from the first 15

minutes of the tilted position, a sensitivity of 81% between the two groups and a specificity of 88% were observed.

After that, the dynamic interactions between cardiac signals (*i.e.* ECG and pressure signal)-which can play an important role in improving the understanding of the cardiovascular regulation - were analyzed. Two different approaches to characterize the relationship between signals in time-frequency domain have been exploited. The first approach is based on the smoothed pseudo Wigner Ville distribution (*SPWD*) and the second uses the local linear correlation coefficient in the time-frequency domain. From these time-frequency spaces, two new indices, able to quantify the relationship between the time series were estimated. The first calculates the density of the coherence between two signals in different frequency bands and the second calculates dot product between the time-frequency spectra of two signals. The advantage of taking into account these interactions is underlined and shown on the same database, an improved performance with a sensitivity of 85% and a specificity of 98% when studying the scalar product in the bandwidth [0.2-0.4 Hz] of the auto time frequency spectrum computed with *SPWD* of the RR-interval and the amplitude of systolic blood pressure time series during the first 15min of HUTT. This study shows also that in high frequency band [0.2-0.4 Hz] the similarity between these two time series is higher in positive than in negative test group.

Although interesting, these findings need to start the tilt test for 15 minutes, which is still unpleasant for the patient in clinical test. Therefore, this work proposes also a method to analyze the dynamics of the heart rate variability in the reconstructed phase space and with hidden Markov models, during the resting phase of the test. The results achieved show a sensitivity of 95% and specificity around 50% using parameters extracted from the histograms in the phase space area.

However, it was also important to assess performance in each analysis window of this study using all recorded modalities (ECG and pressure signals). An Optimal fusion rule is then proposed to combine data from different time series. This combination has led to a significant improvement in performance in different time windows. The Performances reach 100%, 92.8% of sensitivity and 78%, 100% of specificity respectively during the 5 minute of supine position and the first 15 minutes of the tilted position. These results are, to our knowledge, higher than those reported in the literature during the resting phase or early tilted position, and this analysis strategy is becoming the basis of early detection of syncope during HUTT in the clinical domain.

In summary, since the objective of this work was to propose a clinical test of early syncope detection during HUTT, table 7.1 gives a brief comparison between the best performance obtained in each chapter: using linear and non-linear parameters extracted directly from RR-interval time series (chapter 4); using the time-frequency coherence parameters of time series of different features (chapter 5) and using parameters extracted from the transformation of these times series in reconstructed phase space (chapter 6). The results show that

- In different 5-min time windows, the performance is mainly equivalent in chapter 4 and chapter 5 and it is better in chapter 6.
- In Early15 we notice that in chapter 5 when the coherence between different time series is used, the performance is better than when parameters extracted only from one feature time series are taken into account (chapter 4 and 6)
- The best performance is obtained when an optimal fusion rule was applied to combine decisions from different features (*RR-interval, AmpS, dPdt_max, PTT*) in phase space (chapter 6).

TAB. 7.1 – Comparison between the performances obtained in the different chapters.

	Chapter 4		Chapter 5		Chapter 6	
	Se	Sp	Se	Sp	Se	Sp
Supine 5	80%	71%	62.5%	71.4%	100%	86.7%
Early5	74.3%	67.7%	72.1%	64.8%	64.3%	80%
Last5	71.4%	77.4%	77.9%	70.3%	85.7%	93.3%
Early15	88.5%	80.6%	87.5%	98.3%	92.8%	100%

The different studies realized in this thesis have explored many trails to predict the outcomes of HUTT. So the perspectives of this work are multiple:

1. A new database (with fainting and non-fainting subjects) must be acquired in order to confirm and to validate the preliminary results obtained in this PhD.
2. From a clinical point of view, a clinical device working on real time based on the different methods and parameters that have been used must be implemented to help in making decision.

Form a methodological point of view; the findings based on hidden semi-markov models deserve improvements. This can be optimized by exploiting the transformation of the time series in reconstructed phase space (a joint assessment between hidden markov model and reconstructed phase space) but also by integrated into the model different sources with different dynamics. Finally, other cardiovascular signals can be acquired during HUTT and can be exploited using the same methodologies used in this manuscript to evaluate and compare their capabilities to predict syncope during HUTT.

Publications in international journals

[J1] N. KHODOR, G. CARRAULT, D. MATELOT, H. AMOUD, N. VILLE, M. KHALIL, F. CARRE, A. HERNANDEZ. Kernel based support vector machine for the early detection of syncope during head up tilt tests. *Physiological Measurement Journal*. (Published)

[J2] N. KHODOR, G. CARRAULT, D. MATELOT, H. AMOUD, N. VILLE, M. KHALIL, F. CARRE, A. HERNANDEZ. Characterization of the dynamic interaction between cardiac signals using two different approaches for the early detection of syncope during head up tit test. *Digital signal processing journal* (submitted article).

Publications in international conferences

[C1] N. KHODOR, D. MATELOT, G. CARRAULT, H. AMOUD, A. HERNANDEZ, M. KHALIL, F. CARRE. Differentiating between patients with vasovagal syncope using the analysis of heart rate variability during head-up tilt test. *1st International Conference on Communications, Signal Processing, and their Applications (ICCSPA), IEEE sponsor, Sharjah, 12-14 February 2013.*

[C2] N. KHODOR, G. CARRAULT, H. AMOUD, M. KHALIL, P. L'HOSTIS, A. HERNANDEZ. Evaluation of T-wave morphology parameters in drug-induced repolarization abnormalities. *2nd International Conference on Advances in Biomedical Engineering (ICAMBE), IEEE sponsor, Tripoli, Lebanon, 11-13 September 2013.*

[C3] N. KHODOR, G. CARRAULT, H. AMOUD, M. KHALIL, P. L'HOSTIS, A. HERNANDEZ . New T-wave parameters describing repolarization abnormalities induced by drug. *Middle East Conference on Biomedical Engineering (MECBME), IEEE sponsor, Doha, Qatar , 17-20 Feb. 2014*

[C4] N. KHODOR, G. CARRAULT, D. MATELOT, H. AMOUD, N. VILLE, M. KHALIL, F. CARRE, A. HERNANDEZ. A new phase space analysis algorithm for the early detection of syncope during head up tilt tests. *Computing in Cardiology Conference (CinC14), Cambridge, USA, 7-10 Sep. 2014.*

List of figures

1.1 Anatomical structure of the heart.....	7
1.2 Depolarization-repolarization cycle.....	8
1.3 The action potential and the corresponding ionic exchange	9
1.4 The cardiac conduction system	10
1.5 Propagation of the action potential in the heart	11
1.6 Waves, intervals and segments of a heart beat in ECG	11
1.7 The nervous system	14
1.8 General anatomical scheme of ANS	15
1.9 The baroreflex	17
2.1 Classification of brief loss of consciousness	21
2.2 Pathophysiological basis classification of syncope	21
2.3 Another classification of reflex syncope based on (Brignole et al., 2000).....	23
2.4 Example of heart rate and blood pressure evolution in the different types of reflex syncope (DBP, diastolic blood pressure in gray, SBP: systolic blood pressure in black) (Sutton et al 1992).	24
2.5 The Tilt Test	25
2.6 Reproduction of the Bezold-Jarisch reflex at HUTT according to (Kapoor, 1990).....	25
3.1 Data mining framework applied in this manuscript	33
3.2 Example of head-up tilt test during the inclination phase	35
3.3 Example of QRS detection	37
3.4 Left panel: prototype beats corresponding to different morphologies previously detected from the ECG	37
3.5 Alignment and averaging of beats from the same morphology group	37
3.6 Spectral representation of a beat and its components: the P, QRS and T wave	39
3.7 Example of segmentation of the T wave	40
3.8 Definition of the different time series extracted from blood pressure signal	41
3.9 Relief Algorithm	43
3.10 (a) Representative schema of the elliptic search space of SFS, (b) an example of SFS algorithm	44

3.11 An example of ‘probe’ feature selection	45
3.12 Classification for optimal hyperplane	47
3.13 The main ideas of the remaining chapters	48
4.1 Block diagram of the proposed approach for syncope prediction	52
4.2 Characterization plot of different signals measured during HUTT for a negative (upper panel) and a positive (lower panel) subject	53
4.3 Example of an ECG recording with the different time intervals used for analysis, for a subject with positive response	55
4.4 Example of an ECG recording with the different time intervals used for analysis, for a subject with negative response	55
4.5 Example of identification sequences with simultaneous increasing or decreasing in sBP and RR interval	56
4.6 Median and MAE values of parameters with windows of 5 min	59
4.7 Boxplot of the significant parameters in Early15	60
4.8 DFA of RR-interval time series in negative (in red) and positive (in blue) patients examined during the first 15 min of HUTT	61
4.9 The probability distribution functions that the ‘probe’ feature is better classified than a variable candidate of the set of parameters	64
5.1 Block diagram of the proposed approach for syncope prediction	76
5.2 Example of the full time series extracted from ECG and blood pressure of subject with negative response to HUTT	77
5.3 Example of the full time series extracted from ECG and blood pressure of subject experiencing syncope	77
5.4 Different time series extracted from a subject with negative response to HUTT	78
5.5 Different time series extracted from a subject who experienced syncope	78
5.6 The different sub-bands with their numbers as used in this paper	81
5.7 Representation of the RR-interval time series and the <i>AmpS</i> time series and their corresponding TF auto spectrum respectively for a negative and a positive subject during Early15.....	86
B1 Chart representing mode quality by number of components	89

B2 Variable Importance for the Projection	89
B3 ROC curve of the testing phase	90
6.1 An example of trajectory for the x coordinates of a Hénon map	98
6.2 A typical example of a Lorenz system (a) and its recurrence plot (b)	101
6.3 Mutual information representation across different time delays and the percentage of false nearest neighbors across different embedding dimensions.....	103
6.4 RR time series and its reconstructed phase space of for negative and positive subjects with $M=3$ and $\tau=3$	104
6.5 Phase space density plot of (a) negative subject (b) positive subject	105
6.6 Representation of the model quality by number of components	106
6.7 correlation with t1 and t2.....	107
6.8 ROC curve of the testing phase	107
6.9 Optimal fusion rule applied to time series extracted from pressure signal	110
6.10 Difference between HMM / HSMM on the modeling of the time spent in a state..	111
6.11 Time series of the extracted features and illustration of the window in red frame that contains the transition event that happened when the table is tilted	113
C.1 Basic structure of a hidden Markov model	117
C.2 Basic structure of a hidden Markov model	118
C.3 Principle of hidden Markov Model	120
C.4 Principle of hidden semi-Markov Model	12

List of tables

1.1 Characteristics and functions of blood vessels	12
1.2 Effect of sympathetic and parasympathetic activation	16
3.1 Distance functions	46
3.2 kernel functions frequently used in the literature	47
4.1 Median and median absolute deviation values of parameters	60
4.2 Contingency tables in the different windows	62
4.3 Mean performance using 2-fold cross validation with 10 realizations	62
4.4 Results of relief method with 1 is the first and 19 is the last ranking parameter	63
4.5 The average results of 10 iterations of classification using combinations of parameters	64
5.1 Optimal classification results obtained using parameters computed from <i>SPWD</i> coherence	82
5.2 Performance of classification during Early15.....	83
5.3 Matrix representing the average result of the 10 iterations of SFS-KNN during Early15 using S_3 on the test set	83
5.4 Optimal classification results obtained using parameters computed from $R^2(t,f)$	84
6.1 Example of the first 3 cells of a decomposed phase space	99
6.2 RQA parameters	102
6.3 The bounds of the significant cells and their neighbors	104
6.4 Mean and standard deviation of RQA parameters	105
6.5 Matrix representing the average result of the 10 iterations.....	106
6.6 Optimal classification results using parameters from different time series	108
6.7 Optimal fusion results in different time windows	110
6.8 Optimal results obtained by block detection approach	114
6.9 Results of optimal fusion using first only the time series extracted from pressure signal then the time series from ECG and pressure signals	115
C.1 Summary of the main parameters and characteristics of the HMM and HSMM ...	120
7.1 Comparison between the performances obtained in the different chapters	127