Variability monitoring for clinical applications

Andrea Bravi

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements
for the Doctorate in Philosophy degree in Cellular and Molecular Medicine

Department of Cellular and Molecular Medicine
Faculty of Medicine
University of Ottawa

To all who touched my life.

Thank you

Variability monitoring for clinical applications

Andrea Bravi

University of Ottawa

Abstract

Current monitoring tools in the intensive care units focus on displaying physiologically monitored parameters (e.g. vital signs such as heart rate, respiratory rate and blood pressure) at the present moment. Added clinical utility can be found by analyzing how the conditions of a patient evolve with time, and automatically relating that dynamics to population trends. Variability analysis consists of monitoring patterns of variation over intervals in time of physiological signals such as heart rate and respiratory rate. Given that illness has been associated in multiple studies with altered variability, most commonly lack of variation, variability monitoring represents a tool whose contribution at the bedside still needs to be explored. With the long term objective of improving care, this thesis promotes the use of variability analysis through three distinct types of analysis: facing the technical challenges involved with the dimensionality of variability analysis, enhancing the physiological understanding of variability, and showing its utility in real world clinical applications. In particular, the contributions of this thesis include: the review and classification into domains of a large array of measures of variability; the design of system and methods to integrate multiple measures of variability into a unique score, called composite measure, bringing relevant information to specific clinical

problems; the comparison of patterns of heart rate variability during exercise and sepsis development, showing the inability of single measures of variability to discriminate between the two kinds of stressors; the analysis of variability produced from a physiologically-based model of the cardiovascular system, showing that each single measure of variability is an unspecific sensor of the body, thereby promoting multivariate analysis to the only means of understanding the physiology underlying variability; the study of heart rate variability in a population at high risk of sepsis development, showing the ability of variability to predict the occurrence of sepsis more than 48 hours in advance respect to the time of diagnosis of the clinical team; the study of heart and respiratory rate variability in intubated intensive care unit patients, showing how variability can provide a better way of assessing extubation readiness respect to commonly used clinical parameters. Overall, it is hoped that these novel contributions will help promoting bedside applications of variability monitoring to improve patient care.

Table of contents

List of tables

Manuscript 1

Manuscript 3

Manuscript 4

Manuscript 5

Manuscript 6

List of figures

General Introduction

Manuscript 2

Manuscript 3

Manuscript 4

Manuscript 5

Manuscript 6

Acknowledgements

Dato che questa è l'ultima volta (si spera) che avrò l'occasione di ringraziarvi pubblicamente, colgo questa occasione d'oro.

Grazie papà Alberto, e mamma Loredana, perché se sono riuscito ad arrivare a questo punto è grazie a voi. Non è retorica, vi devo davvero tutto, e anche se non colgo mai abbastanze occasioni per ricordarvelo, sappiate che vi sono riconoscente dal profondo del mio cuore.

Grazie anche al mio fratellone Malko, e alla mia sorellina Alessia, perché mi siete sempre vicini, come se vivessimo ancora a Pisa nella stessa camera.

Grazie ad Angela e Gianni per avermi accolto come un figlio, e per avermi donato la cosa più bella della mia vita.

In fine, grazie Giulia, mia sposa... senza di te, come farei? ☺

This section is to say thanks to the two people that made this thesis possible:

Andrew JE Seely
André Longtin

The key for a good PhD is finding good supervisors. I must say that I could not hope for anyone better. You two have guided me and supported me all along these years. There is no way to express how much respect and gratitude I feel towards you. You did a great job.

Also, I would like to thank Geoffrey, the first companion of this long journey: we had so many good talks together! Same applies to Christophe: I can say I never had better lunches together with anyone!

Then I would like to thank the people at Therapeutic Monitoring Systems: Simon, Stephen, Aline, Dave, Warren, Hector… for sharing with me what working in a company means. The beauty of working together to achieve a goal as important as improving care cannot be described.

Lastly, I would like to thank you, the reader. If you got here it is either because you are one of my friends, or because you are interested in how I spent the last three years.

Hope you will enjoy the read!

Chapter 1

1 General Introduction

Data intelligence is everywhere. From the advertisement campaigns of Google and Facebook on the internet, to smartphone applications that monitor position and interests to provide advices on how to spend time and money, data analysis has never been trendier than in these years. The only place where data intelligence still needs to find its place is in the hospitals. This is particularly true for Intensive Care Units (ICUs), where despite the increasing quantity of data recorded and stored, the only intelligence that is present is that of nurses and clinicians. Monitoring tools usually display the current status of a patient, or sometimes a few seconds of data, neglecting the value of trend monitoring. What if the monitoring tools of the future, rather than displaying "raw information" (e.g. the patient has a heart rate of 100 beats per minute), were able to provide clinicians with context-specific, clinically-valuable information? For instance, consider a patient at risk of sepsis development (i.e. systemic infection, in lay terms). A monitoring tool providing an accurate estimate of the probability of developing sepsis at a given time would be preferable to a monitor able to display "raw information" such as heart or respiratory rate. Which data could we use to create such an intelligent system?

Variability analysis consists of monitoring over time patterns of variation within physiological signals. Given that illness has been associated in multiple studies with lack of variation, variability analysis represents a tool whose contribution at the bedside still needs to be explored. This thesis represents an attempt to enhance our ability to utilize variability to create the monitoring devices of the future.

Before starting this journey, it is useful to provide the reader with an overview of the science on which variability analysis is based on. Chapters 1.1 and 1.2 introduce those elements of anatomy and physiology necessary to the understanding of cardiac and respiratory control. After those primers, it will almost seem that the body is a deterministic machine, with a predictable behavior at any given time. The truth is that the body is far more complex, and a paradigm shift is needed to understand its behavior. The complex nature of the body is the focus of Chapter 1.3. Finally, variability analysis is introduced in Chapter 1.4. Chapter 1.5 describes state-of-the-art ICU monitoring systems, bringing the focus back to the clinical problem we are trying to address. To conclude this introduction, Chapter 1.6 summarizes the objectives of this thesis, and provides a summary of its structure, contribution and relevance.

1.1 Fundamentals of cardiac control

This chapter introduces the key concepts for understanding cardiac control. First, the structure of the heart is presented, with a particular emphasis on both its mechanical and its electrical components. Then the cardiac cycle will be described in depth, and with it the electrocardiogram.

1.1.1 Structure of the heart

The heart is an organ whose role is to pump blood across the circulatory system, enabling the simultaneous and continual delivery of oxygenated blood to the body and de-oxygenated blood to the lungs. Circulation accomplishes diffusion of nutrients, thermoregulation, maintenance of homeostasis and more. These functionalities are obtained and maintained by the joint effect of two interrelated components: the **mechanical structure** of the heart and its **electrical conduction system**.

In terms of mechanical structure, the heart is composed of a series of four connected cavities deployed two-by-two, two atria (upper chambers) and two ventricles (lower chambers). Each cavity can be treated as a system with one input, i.e. where the blood enters, and one output, i.e. where the blood is released. In particular, the right atrium receives blood from the **superior vena cava** and moves it to the right ventricle. In between the two cavities is placed the **tricuspid valve**, which avoids blood reflux. The right ventricle then outputs the blood to the pulmonary artery through the **pulmonary valve**. The left atrium receives the blood from the pulmonary vein and moves it to the left ventricle through the **mitral valve** (also called **bicuspid valve**). Lastly the left ventricle

outputs blood to the aorta through the **aortic valve**. The pulmonary and aortic valves are often called **semilunar valves**, while the mitral and tricuspid valves are often called **cuspid valves**. Each of these four cavities is constituted by three layers: the **pericardium** is the most external, it is primarily constituted by connective tissue, which protects the heart by minimizing friction; the **myocardium**, the central layer composed of **cardiomyocytes**, a type of muscle cell that responds to depolarization by contracting; the **endocardium**, primarily made of endothelial cells, which protects the heart and modulates its electrophysiology.

The electrical conduction system of the heart is composed of a set of excitable fibers whose role is to propagate the excitation across the heart, thereby producing the cycle of contractions and relaxations called **cardiac cycle**. The first component of this system is the **sinoatrial (SA) node** (also called **sinus node**), located in the right atrium. The SA node is connected to the **atrioventricular (AV) node**, located in-between atria and ventricles, and the **Bachmann's bundle**, located in the left atrium. The AV node is then connected to **Purkinje fibers** through the **His bundle**. The Purkinje fibers are distributed in both ventricles. A diagram summarizing the structure of the heart is provided in Figure 1.

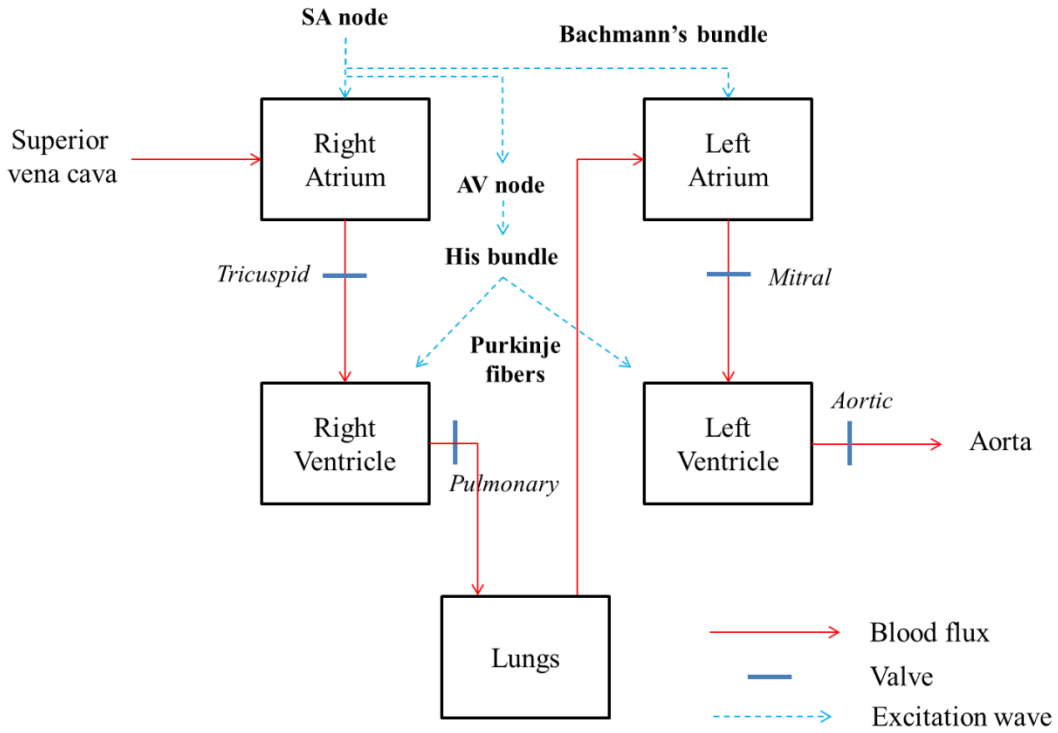


Figure 1 - Schematic diagram of the mechanical structure of the heart and its electrical conduction system.

1.1.2 Cardiac cycle

Now that the mechanical and electrical components of the heart have been listed, it is possible to explain the electrophysiology of the heart, and more specifically how the cardiac cycle is generated. Let's assume we are at the end of the cycle. The ventricles have just contracted, pushing blood in the pulmonary artery and the aorta. This phase of the cycle is called **ventricular systole**. Therefore, the pulmonary and aortic valves are open, while the tricuspid and mitral valves are closed. The ventricle will then relax, producing the closure of the semilunar valves (**ventricular diastole**). The difference in pressure between the atria and the ventricles will cause the opening of tricuspid and mitral valves, starting the accumulation of blood in the ventricles. Then, the electrical

stimulus starts from the pacemaker cells of the SA node, a type of excitable cell able to autonomously generate action potentials. This excitation from the right atrium is rapidly propagated to the left atrium through the Bachmann's bundle, thereby producing the simultaneous contraction of both atria (**atrial systole**). This contraction will conclude the filling of the ventricles. The excitation wave in the meantime is accumulated in the AV node, where it gets propagated through the His bundle to the Purkinje fibers. While that happens the atria relax (**atrial diastole**), producing the closure of the cuspid valves. The excitation through the Purkinje fibers will cause the contractions of the ventricles, after which the cardiac cycle begins anew.

It is important to note that the SA node, AV node and the Purkinje system are all capable of autonomous pacemaker rhythmic oscillation. In normal conditions the SA node sets the pace because its firing rate is higher than the others.

1.1.3 Electrocardiogram

The cardiac cycle we have just described is based on a series of depolarizations of the cardiac cells, leading to consequent contraction. The sum of all these local depolarizations can be globally described as an **electric dipole**, i.e. a physical quantity constituted of two charges of opposites sign separated by a given distance. Therefore, the cardiac cycle can be viewed asthe movement of that dipole in a 3D space. The projections of that dipole along a given plane produces a signal called electrocardiogram (ECG), which has a structure like the one presented in Figure 2. In order: the P-wave is generated by the depolarization of the atria due to the SA node activity, producing atrial systole. The P-Q interval is representative of the propagation of the excitation to the AV node.

When the excitation reaches the His node, the Q-peak is generated. The R-peak is representative of the depolarization moving through the Purkinje fibers and ventricular contraction. The S-peak is representative of ventricular ejection. The T-wave is representative of the repolarization wave moving through the ventricles.

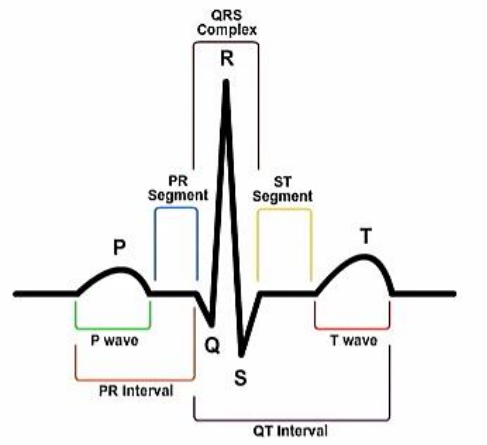


Figure 2 - Sample of electrocardiographic complex[1]

1.1.4 Cardiac control

Cardiac regulation is primarily affected by three components: 1) the autonomic nervous system (ANS), 2) the baroreceptor reflex, 3) respiration[2–4].

The ANS acts through the **sympathetic** and the **parasympathetic** systems. The sympathetic system is distributed over the body through the spinal cord and has a direct effect on multiple organs. In particular, sympathetic stimulation increases the firing rate and the contractility of the heart, constricts blood vessels peripherally and dilates bronchioles in the lung. The heart is stimulated through sympathetic nerves which innervate it entirely and are characterized by a slow action potential velocity (~ 1 m/s). The main synaptic transmitter for the sympathetic nerves is noradrenaline. Systemic

release of adrenaline offers an additional effector arm of the sympathetic nervous system, whose action is mediated by a family of adrenergic receptors. Similarly, the parasympathetic system is located in the spinal cord and affects multiple organs with an antagonistic effect to the sympathetic system. In particular, parasympathetic stimulation decreases the firing rate and the contractility of the heart, dilates blood vessels and restricts bronchioles. This system affects the heart through the **vagus nerve**, which innervates only the SA node and the atria, and is characterized by a high action potential velocity (~ 100 m/s). The main synaptic transmitter for the vagus nerve is acetylcholine. Because this transmitter is more quickly processed at the synaptic level then noradrenaline, and because the vagus nerve has higher action potential velocity, the body responds faster to changes in parasympathetic activity than to changes in sympathetic activity.

The **baroreceptors** are pressure sensors placed in the aorta which respond to an increase in blood pressure by stimulating the aortic and carotid sinus nerves. This stimulation is processed in the **medulla oblongata**, leading to a reduction in sympathetic activity and an enhancement in parasympathetic activity, which will result in a (delayed) decrease in blood pressure.

Respiration influences both blood pressure and heart rate. The mechanisms of autonomic modulation are not fully understood to date[5]. A means by which the heart and lung are coupled is through the alteration in venous return associated with respiration. The **Respiratory Sinus Arrhythmia**[5], which consists of accelerations in heart rate during

inspiration, and decelerations during expiration, allows for the heart to cope with venous return variations.

The aggregated action of ANS, baroreceptor reflex and respiration modulates the cardiovascular function, with the goal of maintaining the **homeostasis** of the body. This is particularly important when the body is stimulated by physiological (e.g. exercise, heat stress) or pathological (e.g. infection) stressors. Depending on the duration and entity of the stressor, we can differentiate the responses as either neurohumoral or proliferative. The **neurohumoral response** addresses stressors on a time scale between seconds and hours, and is produced by the release of chemical mediators (e.g. norepinephrine, angiotensin) into the blood or extracellular medium by the autonomic nervous system. The net effect is the so-called **hemodynamic defense reaction**, which consists of cardiac stimulation, vasoconstriction, and fluid retention. **Proliferative responses** instead are evoked by chronic abnormalities, and produce changes in cell shape, size, and composition through genetic and epigenetic regulation. Further details about the physiology of these responses are out of the scope of the present thesis. For a more detailed analysis we refer the reader to the books of Katz and Opie[6,7].

1.2 Fundamentals of respiratory control

1.2.1 Structure of the lungs

The respiratory system can be divided into airway organs, lungs and respiratory muscles. The **airway organs** are divided between the upper and lower ones. The upper airways include nose, oral cavity and pharynx, and have the role to humidify and warm the inspired air. The lower airways include instead trachea, bronchi and bronchial tree, and have the role to conduct the air to the lungs – i.e. perform **ventilation**. The bronchial tree has a fractal (sees Chapter 1.3) structure, scales down to the bronchioles, which through the alveolar ducts and the alveolar sacs converge into the functional unit of the lungs, the alveoli.

The **lungs** are a collection of hundreds of millions of alveoli, anatomically divided into lobes. In particular, the right lung is divided into upper, medium and lower sections, while the left lung is divided into upper and lower. The lungs are surrounded by double-layered membranes, named pleura. The two pleura create a space called pleural cavity, which enables the lung to expand. Two key parameters of the lungs are elasticity, i.e. the ability to return to their normal shape, and compliance, i.e. the ability to stretch. The diffusion of gas across the alveoli into capillaries and vice versa is called **respiration**.

The **respiratory muscles** include the diaphragm and the rib cage muscles. The diaphragm is the major muscle of ventilation, and its contraction produces an increase in the volume of the thoracic cavity, thereby moving air inside the lungs, a process called

inspiration. The release of air inside the lungs, i.e. **expiration**, is conversely produced by the combined actions of the diaphragm relaxing and elastic recoil of the lungs.

1.2.2 CO_2 capnography

Respiratory signals are monitored in the intensive care unit by recoding the levels of carbon dioxide in the expired air (end-tidal CO_2), called capnography. Capnographic signals, also called capnograms, are commonly acquired using the principle that CO_2 absorbs infrared radiation. Because most of the intensive care unit patients are unable to autonomously sustain breathing, and therefore receive mechanical ventilation, capnograms are easily acquired by attaching the infrared sensor to the tube used for ventilation. An example of capnogram is shown in Figure 3. Clearly, a capnographic signal will increase during expiration and decrease during inspiration.

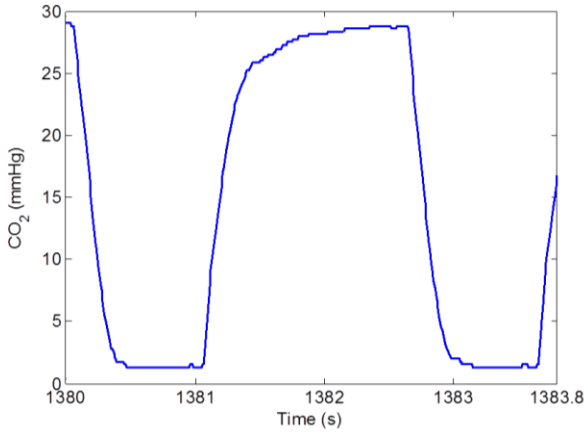


Figure 3 - Example of end-tidal CO2 tracing, called capnogram

1.2.3 Respiratory control

Respiratory control is not understood as well as cardiac control[8]. It is known that the cerebral cortex enables voluntary respiration[8]. However, evidence shows autonomous

respiration is controlled by two oscillatory circuits in the medulla oblongata, one taking care of the timing of inspiration, the other of the timing of expiration. From the medulla the controlling signal is sent in two directions. In one case, the signal passes through respiratory motoneurons in the spinal cord and reaches the diaphragm and the rib cage muscles. In the second case, the signal passes through the cranial motoneurons to glottis, trachea and bronchi, which modulate airway resistance. Further, there would not be a robust control of respiration without feedback mechanisms built in the circuitry. The two primary ones are mechanoreceptors located in the airway, which affect the timing and amplitude of breathing, and chemoreceptors sensitive to pH levels located in the brain, arteries and lungs[8].

1.3 The body as a complex system

So far we have described cardio-respiratory physiology in a deterministic way, almost if the body worked through a set of well-defined rules with no room for randomness. This is contrary to clinical experience. Indeed, it is well known that some patients that seem in bad conditions sometimes manage to recover despite the low expectations, while other patients that seem to be doing just fine sometime deteriorate all of a sudden. Uncertainty is at the heart of clinical practice, and to deal with it there is the need for a paradigm shift: we need to look at the body as a complex system.

A system is a collection of components interacting with each other. When these components interact linearly, the system is said to be a linear system, meaning that when stimulated by any input, the output of the system will always be proportional to that input. Conversely, nonlinear systems present interactions that produce less predictable outputs for given inputs (e.g. discontinuous, saturation, offsets, changes in time scale). Complex systems are nonlinear systems whose behavior cannot be understood by separately inspecting the properties of their components. The nonlinearities within these systems make the principle of superposition invalid, allowing for emergent behavior to arise. Two famous types of emergent behavior that complex systems show are fractality and chaos[9]. The term fractality is used to describe those systems that show self-similarity over time and/or space. In nature there are several examples of fractal structure at different scales, from the fractal folding of the DNA to the fractal structure of interstellar clouds. Chaos, on the other hand, is used to describe systems showing sensitivity to initial conditions and a fractal attractor (i.e. the set of values that N variables defining the state

of an N-dimensional system can assume in an N-dimensional space, namely the phase space of the system). The climate system is an example of complex system showing chaotic dynamics, as demonstrated by Lorentz's model of atmospheric convection[10].

In medical practice the human body is seen as a system showing the following characteristics[11]: 1) it is compartmental, i.e. subsystems of the body (e.g. cardiovascular, nervous, respiratory) can be analyzed separately from each other; 2) it has a dichotomous state, either "normal" or "abnormal", or "healthy" or "diseased"; 3) it has homeostatic dynamics, where clinical parameters are regularly fluctuating around normal values until a perturbation (e.g. a pathogen) affects the body.

It is only in the last 20 years that the human body has started being recognized as a complex system[12–17]. As such, to understand the clinical conditions of a patient, there is the need to simultaneously analyze the body from the point of view of all its subsystems and across different scales, i.e. from genes, omic networks, cells to organs[17–19]. Also, one should allow for more graded transitions between health and illness, and the disease is seen as the transition of the body into a dysfunctional region of its phase and parameter space[11,20]. Lastly, the human body has been reported to exhibit chaotic dynamics and fractal structure[21]. For example, the network of blood vessels and alveoli show a fractal structure which amplifies the diffusion of molecules between tissues. Similarly, the heart rate fluctuates showing self-similar variations over time. The acknowledgment of the complex nature of the human body has led to the development of novel methodologies

for the characterization of the health conditions of a patient. The methodology on which this thesis is based is physiological variability analysis.

1.4 Physiological variability

Physiological variability is commonly associated with the analysis of patterns of variation in clinical "parameters" (using medical jargon) such as heart rate, respiratory rate, blood pressure, percentage of oxygen saturation (SpO_2), temperature and glucose level. The clinical interest in these variations arises from the fact that they convey additional information with respect to the rate averaged over some time window. The earliest examples of this principle are highlighted in a seminal work from the task force of the European Society for Cardiology published in 1996[22], addressing the clinical utility of heart rate variability (HRV). HRV is the study of variations in an R-R interval (RRI) time series, which is the series of time intervals between two successive R peaks in an electrocardiogram (ECG). HRV is the most studied form of variability, possibly due to the simplicity of measuring an ECG, as well as of detecting the position of R peaks. Also, since the Task Force article, several other reviews have assessed the value of performing HRV analysis[23–27]. Applications include risk assessment of mortality after myocardial infarction and cardiac arrest, estimation of heart failure severity, diagnosis of coronary artery disease, diabetic neuropathy, ischemia and sepsis.

The basic principle in variability analysis is studying the properties of an event time series. This is done by taking a certain segment of the time series (also called window) and computing multiple mathematical quantities, called features or measures of variability, each one representing a given property of the time series. Figure 4 depicts how heart rate variability analysis is commonly performed. Starting from an ECG, the

RRI time series is created. Then, a time window of the data is selected. From that window a measure of entropy (i.e. degree of irregularity) is computed.

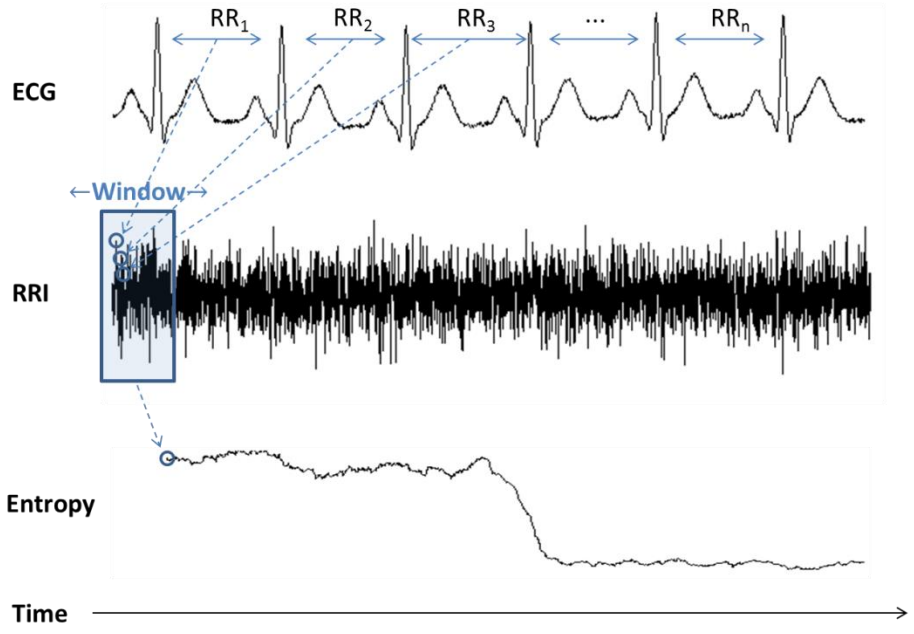


Figure 4 - Example of heart rate variability analysis

The measures of variability are commonly stratified into domains, depending on the type of information they extract. The Task Force in 1996 divided the measures into time, frequential and non-linear. The time domain included statistical features such as mean and standard deviation; the frequential was based on the characterization of the periodicities of the time series through Fourier analysis; lastly, the non-linear domain encompassed any other technique characterizing the scale-invariant properties of the time series, or their degree of complexity. Currently there are more than a hundred measures of variability, and the classification into domains has evolved to adapt to this increasing

number. A focused introduction to the topic, together with a review of several measures of variability with clinical applications, is provided in Chapter 2.

When researchers started focusing on HRV, the underlying hypothesis was that variability provides a window into the autonomic nervous system, and in particular gives insights about sympathetic and vagal activity. Since then, several studies have confirmed the association between frequential characteristic of the R-R interval time series and sympathovagal modulation. In particular, the power at low frequencies is deemed to be related to sympathetic activity, the power at the high frequencies to vagal activity, and the ratio between is therefore an index of sympathovagal modulation[22]. However, with the increasing application of complex system theory to the characterization of clinical parameters, more measures of HRV were developed, and more studies challenged the status quo. Despite its intuitive validity and its long time acceptance, the hypothesis of HRV as a measure of sympathovagal modulation has showed multiple contradictory results[28,29]. For instance, the power at the low frequencies of the heart rate is sometimes associated to sympathetic activity, other times to sympathovagal modulation. Similarly, the power at the high frequencies of the heart rate is often associated to vagal activity, other times to sympathovagal modulation. These results highlighted that this field of research has still much to do to clarify the nature of variability. Although currently there is a lack of consensus about the etiology of variability, additional hypotheses have been proposed. Godin et al. hypothesized that critical illness is associated to decomplexification of the body[20,30]. More specifically, the organ systems are seen as coupled oscillators, and the effect of illness is to reduce that coupling. Morris et al.

extended this view associating uncoupling with reduction in physiological reserve (i.e. increase of physiologic exhaustion)[31]. Seely et al. hypothesized that fractal variability relates to the ability of the body to produce entropy by dissipating energy gradients[32]. This view is inspired by observing natural phenomena such as turbulence, where energy is transferred across waves at different scales, until it dissipates due to viscosity of the turbulent fluid. Doyle et al. analyzed variability from an engineering perspective, hypothesizing that it originates from a simple trade-off between robustness and efficiency, which is common to control systems of any kind[33,34]. In particular, the robustness of a system relates to its ability to maintain its functionality when under stress, while the efficiency relates to its ability to optimally perform its function under ideal conditions. Variability would therefore be the manifestation of adaptability from a higher level controller, i.e. the autonomic nervous system. Therefore, although several hypotheses are available on the nature of variability, there is no shared truth. The advancement of the theory behind variability is limited by the challenges associated with using real-world patient data (e.g. artifacts, nonstationarities), as well as the lack of physiologically-based models capable to reproduce variability.

Besides HRV, other forms of variability have been recently investigated. Blood pressure variability (BPV) is the second most studied type of variability, and is based on the characterization of amplitude variations in systolic and diastolic pressure[35]. BPV has been studied in hypertensive patients to monitor organ damage progression and cardiovascular morbidity[36–39], in particular it has been identified as a risk factor for myocardial infarction[40], coronary heart disease and mortality[41,42]. Respiratory rate variability (RRV)

is the study of variation in the time elapsed between two successive breaths. RRV has been primarily applied to assess the ability of mechanically ventilated patients to breath autonomously[43–45]; other pilot applications are the monitoring of the effects of sedation in intensive care unit patients[46] and the detection of upcoming illnesses in neonates[47]. Temperature variability (TV) is the study of fluctuations in temperature levels of specific areas of the body[48,49]. TV relates to aging[50] and severity of organ failure[51,52], and can identify sepsis development[53].

Although there is a lack of understanding of the physiology underlying variability, its analysis provides novel insights into the dynamics of clinical parameters. The body being a complex system, we know that clinical parameters are far from being stable or strictly periodic, and therefore their continuous monitoring is needed to know the conditions of a patient. Luckily enough, technological development has currently opened up novel possibilities for patient monitoring.

1.5 Monitoring systems in Intensive Care Units

Technology is pervasive in ICUs. Clinicians are provided with multiple tools to assess the conditions of a patient at a given time. Furthermore, continuous monitoring is nowadays routine for multiple physiological variables. These improvements bring a massive amount of information to the clinicians, without however saying anything on the actual conditions of a patient. The future in ICU monitoring is therefore to translate "raw information" into scores to support clinical decision. In particular, these clinical decision support systems (CDSS) are characterized by three features: 1) use of historic trends; 2) confrontation with population data (database storage); 3) integration into a clinically meaningful score. The usage of historic data from a patient allows characterizing the dynamics (i.e. change over time) of physiological parameters of interest. This approach overcomes a current limitation in patient monitoring, which is the focus on the present conditions of a patient, by giving the ability to identify early signs of deterioration[20]. The second feature of CDSS is to assess the conditions of a patient confronting his status with the ones of a population of similar cases. This approach accounts for the large between- and within-subject variability, common in clinical data. Through the application of statistical models, this variability is put into context, making the assessment of the patient's conditions more objective. The third feature of CDSS is the integration of multiple pieces of information into a metric that is directly related to the clinical problem of interest. In this way, the clinician is released from the overburden of information, and only what is clinically useful is delivered, providing added value to his clinical decisions.

To date there are only few systems that embodied these principles. A first example is Visensia™ (formerly Biosign™), from OBS Medical. This monitoring system combines heart rate, respiratory rate, blood pressure, temperature, and oxygen saturation into a unique score – the Visensia index - which tracks the deviation of the status of the patient from "normality" (built from a database of cases), and is then used for the early warning of patient deterioration[54,55]. Despite the promising preliminary results, the system did not reduce mortality or number of adverse events in a randomized control trial (RCT) of 402 high-risk medical and surgical patients[56]. The reason for this failure seemed not to be in the decision model, but rather in how it was used at the bedside. For instance, although a retrospective analysis showed that the monitors were properly displaying abnormalities in physiological variables, no action occurred possibly because of an insufficient number of nurses observing the monitors. Also, the identification of an abnormality does not bring information about how to treat the patient, thereby not producing an improvement in statistics such as mortality or number of adverse events. A second example is the Hero™ monitoring system, from Medical Predictive Science Corporation. This CDSS is used to predict neonatal sepsis in Neonatal Intensive Care Units (NICU) [57]. In a RCT of 3000 low birth weight infants (less than 1500g), the Hero monitor reduced mortality by 2%. The reduction in mortality doubled when the analysis was focused only on extremely low birth weight infants (less than 1000g)[58].

1.6 Thesis overview

1.6.1 Objectives

The present thesis has three objectives. In the first place, we want to characterize and improve the methodology behind variability analysis. The second objective is to deepen the physiological meaning of variability. Lastly, we wish to focus on the creation of decision support systems for specific clinical applications based on variability. The underlying hypothesis is that variability can provide added value to standard clinical practice. The ultimate goal of this research is to design methods and tools to improve care.

1.6.2 Structure, contributions and relevance

The next chapters will present the scientific articles published in peer-reviewed journals, as well as a non-provisional US patent, which were generated during the PhD. Each chapter will present a commentary section, introducing the research reported in the contribution section, where the manuscript is reproduced as published.

In Chapter 2 a comprehensive array of measures of variability is introduced, highlighting their rationale and mathematical basis, and providing references of clinical applications. In this section technical challenges of variability analysis are discussed, and a novel classification of the measures into domains is presented. The investigation of this wide array of measures of variability filled a gap present in the literature, and was justified by the assumption that a heterogeneous set of metrics is required to demonstrate the utility of variability in clinical applications of different nature.

Chapter 3 delves into the importance of merging the relevant information coming from multiple measures into a clinically-relevant and problem-specific index, which the clinicians can use to improve care. Data overload is a common problem in the ICU, and methods such as the ones shown in the presented patent are necessary to translate variability analysis into a clinically usable tool.

Chapter 4 investigates the meaning of HRV measures using a physiologically-based model, and validates the generated hypotheses on real patient data. This analysis shows that each measure of variability is an unspecific sensor of the body, and that in order to gain understanding of the physiology underlying variability a multivariate approach is needed.

Chapter 5 presents two clinical decision support systems. The former is a monitoring system associated with an alarm for the identification of sepsis development, which showed the ability to produce correct alarms more than 48 hours in advance of sepsis diagnosis from the clinicians. The latter is a system predicting extubation outcome, which showed increased sensitivity and specificity with respect to commonly used clinical variables.

Chapter 2

2 Methods for variability analysis

2.1 Commentary

The concept in variability analysis is studying the properties of an event time series. This is done by taking a certain segment of the time series and computing multiple mathematical quantities, called features or measures of variability, each one representing a given property of the time series. This seemingly simple process hides a variety of technical challenges, which can be grouped in: 1) parameter selection and 2) compliance to background hypothesis. The first challenge refers to the fact that variability analysis is based on a wide array of parameters. Each measure of variability requires at least one parameter to be fixed, i.e. the length of the time series to analyse, as well as measure-specific parameters, such as the number of bins to compute the Shannon entropy or the embedding dimension and delays for the recurrence plot measures. Those measures are better explained in the article associated to this chapter (see 2.2). The second challenge refers to the fact that most of the measures make a variety of assumptions: some assume stationarity of the event time series, others the absence of noise. Whether the application of a measure when those assumptions are not respected produces a reliable result or not is an open question. This is particularly true for possibly the most critical problem, which is the assumption that computing the value of a measure for a given time series makes sense. Although banal, this is not straightforward when the underlying nature of the event time series under study is unknown. For instance, the nature of the R-R interval time series (i.e. the series of time elapsed between two R peaks in an electrocardiogram) has

been long debated, with some researchers promoting the idea of it being chaotic, and others rather seeing it as purely stochastic[59]. Far from having a unique answer to this problem, it is clear that measures of chaoticity in a deterministic system, such as the largest Lyapunov exponent (measure of exponential divergence of trajectories in the phase space), may well lead to inconsistent results when computed for a stochastic time series[59].

This impasse cannot be solved without a deep understanding of the etiology of variability, likely aided by the creation of models able to quantitatively reproduce physiological variables, both in healthy and diseased conditions. Whether the creation of such models is feasible or not is again a subject of intense research. More pragmatically, researchers focusing on clinical problems rely on a black-box approach, deciding to use a given set of measures not because their physiological meaning is known, but rather because they have been proven useful in previous studies, or simply because they are supposed to provide independent pieces of information. Which measures should be chosen for a given clinical problem, and how many, are open questions. Researchers started stratifying the measures in different domains exactly to address the first question. In particular, each domain of variability represents the type of information that a measure from that domain extracts. Because several mathematical quantities have multiple interpretations (e.g. the standard deviation of a signal can be used to define its distribution, if the signal is a realization of Gaussian noise, or more generically can represent its energy), there is no strong boundary between the domains. However, the classification of measures into different domains allows to a priori select measures that

are deemed to track different properties of an event time series. The ultimate hope is that among those properties, there will be one or more related to the clinical outcome of interest (e.g. prediction of sepsis development, early warning of deterioration, etc…).

To aid the selection of appropriate measures, we performed a literature review on variability measures. The following article presents an up-to-date list of measures for variability analysis, providing for each an intuitive description, examples of clinical applications, technical references, and a novel classification into domains.

2.2 Manuscript 1: Review and classification of variability analysis techniques with clinical applications

Published in Biomedical Enginnering Online in date 10 October 2011.

My contribution: literature review, writing of the manuscript, creation of the novel classification into domains, review and final approval.

Review and Classification of variability analysis techniques for clinical applications

Andrea Bravi[1*], André Longtin[2], Andrew J.E. Seely[1,3,4]

[1]Department of Cellular and Molecular Medicine, University of Ottawa,
[2]Department of Physics, University of Ottawa
[3]Division of Thoracic Surgery, University of Ottawa
[4]Department of Critical Care Medicine, University of Ottawa

Abstract

Analysis of patterns of variation of time-series, termed variability analysis, represents a rapidly evolving discipline with increasing applications in different fields of science. In medicine and in particular critical care, efforts have focussed on evaluating the clinical utility of variability. However, the growth and complexity of techniques applicable to this field have made interpretation and understanding of variability more challenging. Our objectives are to provide an updated review and evaluation of variability analysis techniques suitable for clinical applications. We review more than 70 variability techniques, providing for each technique a brief description of the underlying theory and assumptions, together with a summary of clinical applications. We propose a revised classification for the domains of variability techniques, which include Statistical, Geometric, Energetic, Informational, and Invariant. We discuss the process of calculation, often necessitating a mathematical transform of the time-series. Our aims are to clarify a broad literature, promote a shared vocabulary that would improve the exchange of ideas and the analyses of the results between different studies. We conclude with challenges for the evolving science of variability analysis.

Introduction

Variability analysis can be defined as the assessment within a system of the degree and character of patterns of variation over time intervals. This analysis found applications in many different research fields, from weather forecasting [60], to network analysis [61], process monitoring [13] and obviously medicine, the subject of this paper. Considering the systemic host response to trauma, shock, or sepsis as a complex system, Seely et al. hypothesized that the analysis of multiorgan variability (patterns of variation over time) and connectivity (patterns of interconnection over space) of physiological signals offer a means to track the system state of the host response, offering the potential for early detection of deterioration and improved real-time prognostication [13,62]. Rooted in non-linear dynamics and physics, the approach of variability analysis has been successfully applied also for the prediction of mortality after acute myocardial infarction [63,64], detection of sleep apnea [65,66], assessment of the autonomic nervous system activity [67,68] and evaluation of the circadian rhythms regulating the body [69,70]. In particular, there is an increasing interest in the application of variability monitoring to improve clinical outcomes [15].

There has been extensive research to develop multiple measures of variability and trying to characterize how variability changes with respect to specific stimuli. Our objective in this paper is to take a step back and assess the array of variability techniques available today, analyze their properties and clinical applications, and evaluate their classification. The most specific classification sees several *domains of variability* such as the Time domain, the Frequency domain, the Entropy domain and the Scale-Invariant domain [15].

However, in many recent papers the classification is often reduced to only Time domain, Frequency domain and a more general Nonlinear domain [67,71–74] . Therefore, despite the several years passed, the classification of variability techniques has not significantly evolved since the important work of the Task Force for Heart rate variability [22].

The objective of the present paper is to present the variability techniques currently available, with an emphasis on the ones used for the analysis of physiological signals in clinical settings. Furthermore, a new classification of these techniques is proposed with the aim of categorizing the knowledge collected in different fields using variability, giving the possibility to new researchers approaching variability analysis to orient themselves in the analysis of their results, and aiding the exchange of ideas. The paper starts with the explanation of the proposed classification, useful for the comprehension of the variability techniques. Then, each technique is presented, highlighting its assumptions, giving an intuitive explanation of the mathematical procedure to compute it, along with examples of application in clinical contexts. The paper concludes with the presentation of two major challenges in variability analysis and the conclusions. In the following, "time series" is meant as a set of measurements of one variable. When the context is the simultaneous measurement of multiple variables, we will use the term multivariate time series.

Classification

After the work of the Task Force of 1996 [22], there has been considerable development of techniques for variability analysis. We highlight two main categories of variability

techniques: *Transformations* and *Features*. Taken from both linear and nonlinear analysis, those techniques share in common the potential to manipulate data in order to identify certain properties of interest. Starting from the former, we here define a *transformation* as mathematicians would do, i.e. as a function which maps the samples of dataset from one set of values to another. Examples of transformation techniques are Recurrence Plots [75,76], Poincaré Plots [67] and Symbolic Dynamics [77]; others will be shown in the following. The principle underlying these techniques is that they rearrange data in ways allowing the extraction of additional features, otherwise difficult to detect. The payoff is that the enhancement of a certain type of information weakens the detection of other types of information. Several transformations have been devised by a range of scientists for time series analysis; we categorized them in two different sets due to their conceptual difference. The first one is what we call the *Quantitative Transformations*. All the techniques belonging to this set make use of specific mathematical models to reshape a time series. The second one is the set of *Qualitative Transformations*, whose aim is to quantize the dataset, without imposing particular models on the time series. The potential advantage of the quantization, which compresses a certain range of values of the data to a specific value, is that it helps avoiding the problems associated with the noise in the data, and it can make the analysis more robust to artifacts. However, the clear result is the loss of potentially valuable information.

The second category of variability techniques is what we called *features*. This term is commonly used in signal processing to identify a specific piece of information (feature) that can be extracted from the potentially available information in the signal. The aim of

variability analysis is to extract a number of features from the data, and if needed (and possible given the amount and quality of the data) to track their temporal evolution. In many cases, it is required to apply a transformation to the data before feature extraction, as this enables the extraction of features otherwise concealed or inaccessible. For example, the Fourier transform is a transformation, and from it one can extract multiple features like the total power of the signal, the power at specific frequencies and the power within specific frequency ranges. When features are analyzed, it is useful to categorize them depending on the types of informative content they are extracting from a system, namely, what in the literature are called *Domains of variability* [15,71–74]. The most specific classification sees several domains, such as the Time domain, the Frequency domain, the Entropy domain and the Scale-Invariant domain [15]. However, this classification is not enough to cover all the techniques of variability analysis available today. Our review has identified five different domains of variability: **Statistical**, assuming the data are from a stochastic process, the associated features describe the properties of their distribution; **Geometric**, which describes those properties that are related to the shape of the dataset in a certain space; **Energetic**, describing the features related to the energy or the power of the data; **Informational**, describing the degree of irregularity/complexity of a system; **Invariant**, describing the properties of a system that are not supposed to change over either time or space. The chart of the domains, together with the list of the features analyzed in this review, is reported in Table 1.

Table 1 - Description of the domains characterizing the features

Types of domain	Name
Statistical	Form factor, Lee parameter, some Symbolic dynamics features, Standard statistical features, Turns count.
Geometric	Grid counting, Dynamical moments, Heart rate turbulence, Poincaré plots features, Recurrence plot features, Spatial filling index.
Informational	Approximate entropy, Conditional entropy, Compression entropy, Fuzzy entropy, Kullback-Leibler permutation entropy, Multiscale entropy, Predictive-based features, Sample entropy, Shannon entropy, Similarity indexes, Rényi entropy.
Energetic	Energy operators, Frequency features, Multiscale time irreversibility, Time-Frequency features.
Invariant	Allan scaling exponent, Correlation dimension, Detrended fluctuation analysis, Diffusion Entropy, Embedding scaling exponent, Fano scaling exponent, Finite growth rates, Higuchi's algorithm, Index of variability, Kolmogorov-Sinai entropy, Largest Lyapunov exponent, Multifractal exponents, Power spectrum scaling exponent, Rescaled detrended range analysis, Scaled windowed variance.

The link with the previous classification [15] is that Time domain is now divided into Statistical and Geometric; the Frequency domain got re-assigned to a more general Energetic domain; the Entropy and the Scale-Invariant domains have become, respectively, the Informational and part of the Invariant domain.

We will now analyze the different type of transformations and domains, describing in depth the assumptions underlying each technique, discussing benefits and limitations. The list of all the analyzed variability techniques in alphabetic order can be found in Table 2.

Table 2 - List of the techniques for variability analysis

Type	Name	Clinical References	Technical References
Transformations	Angle map	[78]	[78]
	Blackman-Tuckey's method	[79]	[80]
	Burg's method	[81–83]	[80]
	Grid transformation	[84]	[84]
	High order spectra	[85,86]	[85,86]
	Integral pulse frequency modulation	[87,88]	[87]
	Lomb's method	[83,87,89]	[89]
	Multitaper	[90,91]	[90,91]
	Phase space representation	[59]	[10,92]
	Poincaré plots	[67]	[67]
	Point process	[93–95]	[93–95]
	Recurrence plots		[96]
	Rhythmometric analysis	[71,97]	[71,97]
	Short time Fourier transform	[98]	[98]
	Symbolic dynamics	[82,99–103]	[82,99–103]
	Wavelet transform	[90]	[104]
	Welch's method	[83]	[80]
	Wigner-Ville transform	[105]	
	Yule-Walker's method	[83,99]	[80]
Features	Allan factor	[95,106,107]	[95,106]
	Approximate entropy	[62,108–110]	[111]
	Compression entropy	[74,112]	[112]
	Conditional entropy	[82]	[82]
	Correlation dimension	[113]	[92]
	Detrended fluctuation analysis	[62,114,115]	[114–116]
	Diffusion entropy	[102,117,118]	[119]
	Dynamical moments		[120]
	Embedding scaling exponent	[121]	[121]
	Energy operators	[122,123]	[122,123]
	Fano factor	[95,106,107]	[95,106]
	Finite growth rates	[103]	[103]
	Form factor	[124]	[124]
	Fuzzy entropy		[111]
	Frequency features	[22,25,67,79,82,125]	[22,25]
	Grid counting	[84,126]	[84,126]
	Heart rate turbulence	[127]	[127]
	High order spectra features	[85,86]	[85,86]
	Higuchi's algorithm	[102,128]	[128]
	Index of variability		[61]
	Kolmogorov-Sinai entropy	[113]	[92]
	Kullback-Leibler permutation entropy	[129]	[129]
	Largest Lyapunov exponent	[130,131]	[10,92]
	Lee parameter		[132]
	Multifractal exponents	[102,133,134]	[133,134]

Multiscale entropy	[135]	[135,136]
Multiscale time irreversibility	[136]	[136]
Permutation entropy	[129]	[129]
Poincaré plots features	[62,67,102,125]	[67,102]
Power spectrum scaling exponent		[116]
Predictive-based features	[137–139]	[137–139]
Recurrence plot features	[75,96,140–142]	[96]
Rényi entropy		[82]
Rescaled detrended range analysis		[116]
Sample entropy	[62,108–110]	[111]
Scaled windowed variance		[116]
Shannon entropy	[100,101,103,138]	[82,138]
Similarity indexes	[84,99,143,144]	[84,99,143,144]
Symbolic dynamic features	[82,99–103]	[82,99–103]
Spatial filling index	[145]	[145]
Standard statistical features	[22,67,74,82,102,146,147]	[22]
Time-Frequency features	[22,25,67,79,82,125]	[22,25]
Turns count	[124]	[124]

Transformations

Quantitative

Power spectrum

The *power spectrum* is a method that assumes the signal under study comes from a stationary stochastic process. There are several methods to extract the power spectrum of a signal (i.e. *Frequency transformations*). Among the ones used in clinical settings are the *Blackman-Tuckey's method* [79], the *Welch's method* [148], the *Burg's method* [81,82], *Yule-Walker's method* [99], the *Lomb's method* [89] and the *Multitaper transform* [90,91]. The performance of these methods in the extraction of the power spectrum depends on the nature of the dataset under study, and its study represents a research field *per se* [83,87]. In general, it is not possible to select one technique to suit any particular application. Just to provide an example, above the ones cited, the Lomb method is the only one designed to create the power spectrum of an unevenly sampled time series; as such, it is preferable in

the case of RR interval analysis, whenever interpolation should be avoided [89]. The above being classical signal processing techniques that have been extensively studied, their detailed analysis would be out of scope in this context; we suggest the interested reader to refer to [80] for a detailed explanation and comparison of their properties. Similar considerations can be made when *higher order spectra* (HOS) are studied. To explain in simple words what HOS are, it must be recalled that the power spectrum of a signal is equivalent to the Fourier transform of its autocorrelation function, i.e. the second order moment of the signal (via the Wiener–Khintchin theorem). Whenever the Fourier transform of moments higher than the second one are desired, a higher order spectrum can be obtained. HOS are particularly suited for the analysis of nonlinear systems and non-Gaussian signals [85,86].

Time-Frequency

All the power spectrum estimation techniques presented so far are based on the assumption of absence of nonstationarities in the studied signals, implying that their properties do not change with time. In clinical settings this is not always the case, even if this rule is often neglected trying to study short-time recordings. To extend this analysis to nonstationary signals different techniques were introduced (i.e. *Time-Frequency transformations*). The most classical approach involves a *windowed analysis* of a time series, i.e. to the division of the time series in multiple segments (also said windows). For every window the power spectrum is computed using the Frequency transformations just cited. A standard is the *Short-time Fourier Transform* (STFT) [98], where the time-frequency transformation is computed just using the Fourier transform of the considered window. Because those methods relying on a windowed analysis suffer of the trade-off

between frequential and temporal resolution (large windows give high frequential resolution but low temporal resolution, and vice versa), another class of techniques was introduced. The *Wigner-Ville transform* (WVT) [105] is a technique based on the Fourier transform of the instantaneous autocorrelation of the signal, and the *Wavelet transform* (WT) [104] is based on the correlation at different scales and time between the data and a reference signal, called mother Wavelet. The former is capable to give the finest temporal and frequential resolution compare to the other techniques, but is highly sensitive to noise; the second was developed to give high frequential resolution at low frequencies and high temporal resolution at high frequencies. Time-Frequency transformations are classical signal processing techniques with a broad literature behind; therefore, their detailed discussion is out of scope.

Integral pulse frequency modulation

The *Integral pulse frequency modulation* (IPFM) is a model hypothesizing that the sympathetic and parasympathetic systems influences on the sino-atrial node can be associated to a single modulating signal, which integral generates a beat every time a threshold is surpassed [87]. Given a RR interval time series, from this model it is possible to extract the spectrum of the modulating signal, and from it multiple Energetic features. The benefit of the IPFM model is a direct physiological explanation of the variability of a RR interval time series; the limitations consists in neglecting the effect of other systems (e.g. respiration, hormonal activity), and imposing a simple model describing the autonomic nervous system regulation of the heartbeat. The features that can be extracted from the IPFM model correspond to all the ones that will be presented for the power spectrum. This model was recently extended to enhance the computing of Heart Rate

Turbulence [88], a Geometric feature successfully employed to assess the risk of myocardial infarction.

Phase space representation

The *Phase space representation* is a transformation mapping a time series into a multidimensional space, where each dimension represents an independent variable describing the system under study. There are several variants of this transformation [92], however the most famous is extracted from Takens' embedding theorem [10]. Taken's theorem justifies the transformation of a time series into a **m**-dimensional multivariate time series. This is done by associating to each **m** successive samples distant a certain number $\boldsymbol{\tau}$ of samples, a point in the phase space. This technique is used for the characterization of dynamical systems, and it is mathematically proved that it can be used to recreate the attractor of the dynamics characterizing the signal (an attractor is a set of states in which the dynamics converge over time). The main limitation is that the attractor can be reconstructed on the assumptions that 1) the time series object of study is without noise and 2) infinitely sampled. Despite this is not the case for clinical applications, its usefulness is well assessed in the literature. The reason is that the Phase space representation is mathematically equivalent to a multidimensional representation of the autocorrelation of the signal, making it useful also when the degree of determinism of a time series is not clear (e.g. heart rate variability [59]). This transformation is the starting point of other types of transformation, like Recurrence plots, Poincaré plots, Grid transformation and certain types of Symbolic dynamics.

Recurrence plots

Recurrence plots are a visual representation of all the possible distances between the points in the phase space [96]. Whenever the distance between two points is below a certain threshold, there is a recurrence in the dynamics: i.e. the dynamical system visited multiple times a certain area of the phase. From this transformation, well suited for the study of short nonstationary signals, many Geometric features can be extracted.

Poincaré plots

The *Poincaré plots* are a particular case of phase space representation created selecting **m=2** and **τ=1**; that corresponds to displaying a generic sample **n** of the time series as a function of the sample **n-1** [67]. This is also known as a *return map*. The main limitation of this technique is that assumes that a low dimensional representation of a dynamical attractor is enough to detect relevant features of the dynamics. Despite its simplicity, this transformation was successfully employed also with high dimensional systems like the Heart Rate variability. The benefit is that, given the low dimensionality, it is possible to easily design and visualize several types of Geometric features. A variant of the Poincaré plot is called the *Angle map*, which is a Poincaré plots made after that the transformation of the data into a polar coordinate system [78].

Grid transformation

The *Grid transformation* corresponds to the representation of the data into a bi-dimensional phase space (i.e. **m=2**, with tunable embedding delay **τ**) that is divided into a fixed number of squares called pixels. This process creates a binary image of the phase space, where 0 is assigned to the pixels that were not visited by the dynamics, and 1 to the pixels that were visited [84].

Rhythmometric analysis

The *rhythmometric analysis* (RA) models the changes inside a time series as a set of cosine functions with linked frequencies [71,97]. Three are the main components of this kind of model: circadian rhythm (with a given frequency *f*, considered as the fundamental frequency of the time series), infradian rhythms (which frequencies are below *f*) and ultradian rhythms (which frequencies are multiples of *f*). Therefore, RA is a frequential analysis that tries to explain the changes at different time scales of the data. The main difference with standard frequential analysis is that only particular frequencies are taken into account, and usually there are additional terms in the model that take into account the variability of the frequencies of the components inside the model. This transformation can be employed to explain the changes of features over time, but also to create a purely deterministic version of the original time series that can be studied more easily because of the absence of noise [71,97].

Point processes

Point processes are a family of stochastic models modeling the time intervals between successions of events. In this kind of model the information is held only in the intervals between the events, and not in the event *per se*. An RR interval time series by definition can be seen as a point process [93–95]. As any stochastic model, a certain probability distribution, in the case of point processes called *Conditional Intensity Function*, is used to draw at which time the next event (i.e. a beat) will take place. To follow the nonstationarity of the signal, the parameters of the model can be estimated through adaptive filtering [93–95].

Qualitative

Bin transformation

In the case of a *bin* transformation, the range of the time series (determined by the minimal and maximal value of the variable) is divided in a specific number of equally spaced intervals of arbitrary size. Every time a data point falls in a given interval, the count in the associated bin is increased by one. This transformation is the one commonly employed in the construction of histograms.

Symbolic dynamics

The purpose of *Symbolic dynamics* is to transform a time series into a sequence of specific symbols. Plenty of possible approaches can be used to coarse-grain a time series, however the basic idea is always to divide the range of the time series into intervals and assign to each interval a symbol. This process is usually employed either considering the signal as a whole or considering windows of the time series. When this process is separately applied to windows of the time series, the absolute information held by the signal is lost (i.e. detrend of the time series), keeping only the information related to its shape. Symbolic dynamics transformation is often applied in combination with the Phase space representation; this approach corresponds to the division of the phase space into a set of hypercubes [82].

Being only a way to discretize the dataset, the symbolic dynamics transformation is associated to features belonging to different domains of variability. A set of features contains Informational domain measures, usually applied also with the bin transformation (e.g. Shannon entropy, Rényi entropy, Similarity based indexes). The other set contains

statistical features that belong only to the symbolic dynamics. Each set of features will be discussed in the appropriate paragraph below.

Features

Statistical

The Statistical domain assumes the data as a stochastic process; the following features describe the properties of their distribution. For a summary of the key concepts, one can refer to Table 3.

Form factor

This feature is based on the standard deviations of the signal and its first and second derivative (a derivative represents an infinitesimal change in the signal). The computation involves the ratio between 1) the ratio of the standard deviations of the second and the first derivative of the signals, and 2) the ratio of the standard deviations of the first derivative and of the signal. The Form factor was found useful for the characterization of Vibroarthrographic signals (sounds coming from the movement of the knee-joint, describing the mechanical properties of articular cartilage surfaces) [124].

Lee parameter

This is used to assess the best window size to compute variability analysis in telecommunication [132]. It quantifies the dispersion of the data by the ratio between 1) the mean value plus the standard deviation, and 2) the mean value minus the standard deviation of the time series. The more similar the two values are, the higher is the Lee parameter**.**

Symbolic dynamics features

Those are based on the transformation of a time series into a sequence of symbols through the association of a range of value with a specific symbol. There are two types of features characteristic of symbolic dynamics:

- *Forbidden words*, which count the number of words (a succession of symbols) with a given length that do not appear. This technique was applied for the evaluation of heart rate and blood pressure variability in patients with dilated cardiomyopathy [100].
- *Percentage of variations*, which quantifies the number of variations inside the words composing a time series. For example, if words of 3 symbols are under study, the word **[a a a]** has no variation (all symbols are identical), the word **[b b a]** has one variation (two successive symbols are identical), the word **[a b a]** has two variations (there are no successive repetitions of symbols). This technique was applied in the characterization of autonomic cardiac modulation during arrhythmias [101] and the characterization of chronic heart failure [102].

Other features used with symbolic dynamics transformations can be found in the Informational domain (such as Shannon entropy, see below).

Standard statistical features

Those features were previously collected by the Task Force in 1996 to study heart rate variability [22]. They involve mean, standard deviation, counting of the samples above or below a certain threshold, and other statistical measures that are based on the distribution of the data and not their order (the list can be found in Table 5).

Table 3 - List of Standard statistical features

Short name	Description
SDNN	Standard deviation of the Normal-Normal (NN) interval.
RMSSD	Square root of the mean squared differences between each successive NN interval and the mean interval.
pNN x	The ratio between the number of successive NN intervals that are greater than x and the total number of NN intervals. Usually x is set to 20 ms.
NN x	It is the number of successive NN intervals that are greater than x. Usually x is set to 20 ms.
HTI	HRV Triangular Index: ratio of the total number of RR intervals to the number of intervals in the bin with the highest number of intervals The bin width should be around 1/128 seconds (128 Hz is the standard).
TINN	The triangular interpolation of the NN interval histogram (TINN) is the width of the base of the triangle that best approximates the NN interval distribution (the minimum square difference is used to find such a triangle).
IQRNN	Interquartile range of NN. It is the difference between the upper and the lower quartile values of the NN interval distribution (25% of the samples above and below the median value) [54].
SDSD	Standard deviation of the first derivative of the time series [138].

To limit the complexity of the analysis, in the case of heart rate variability those measures are usually applied to the normal R-normal R time series, where the prefix "normal" indicates the R peaks that are not classified as premature (i.e. ectopic beats) or extracted from known cardiac electrical dysfunctions (e.g. fibrillation). Despite their simplicity, the value of this measure has been proven in innumerable clinical applications like the prediction of sepsis [149], the assessment of the risk of myocardial infarction and diabetic neuropathies [22,25], the assessment of cardiac contractility [125], and the classification of cardiac death [82], chronic heart failure [74], and severity of Parkinson's disease [67].

Turns count

Turns count records the number of times that the signal is above or below a certain threshold. This measure can be applied successfully only when specific properties of the signals are investigated. Following this reasoning, the threshold can be either fixed or

adaptive. An adaptive threshold using the standard deviation of the analyzed signal was used to classify vibroarthrographic data from different knee-joint pathologies [124].

Geometric

The Geometric domain describes the properties that are related to the shape of the dataset in a certain space. For a summary of the key concepts refer to Table 6.

Grid counting

This feature estimates how much a bi-dimensional attractor fills its phase space, and is based on the Grid transformation. Essentially, the Grid counting is the number of pixels visited by the dynamics divided by the number of pixels in the grid. The Grid counting was applied to the detection of ventricular fibrillation [84,126].

Dynamical moments

Dynamical moments are a set of features that statistically describe the geometrical properties of the phase space representation of a time series. In particular, the moments of second order and above are calculated assuming each dimension of the phase space is a random variable. The dynamical moments have been applied to the analysis of signals from chemical sensors [120].

Heart rate turbulence

Heart rate turbulence is a technique applicable only to heart rate variability. It is based on the evaluation of ventricular premature complexes (VPCs), and studies how VPCs modify the pattern of an RR interval time series. In normal subjects, the heart rate following a VPC incurs an early acceleration followed by a late deceleration. Heart rate turbulence extracts indexes describing those two phenomena through, respectively, the *Turbulence onset* and the *Turbulence slope*. The Turbulence onset is given by the

percentage difference between the two RR intervals preceding and following a VPC, and is an index of the early deceleration. The Turbulence slope is given by the maximum positive regression slope assessed over any 5 consecutive beats within the first 15 beats after the VPC, and is an index of late deceleration. The sequence of RR intervals before and after a VPC is called VPC tachogram. Given the variability between the heart rate pre- and post-VPCs, Turbulence onset and Turbulence slope are measured over the so-called average VPC tachogram, i.e. the average of the VPC tachograms over usually 24 hours (at least 5 VPCs are needed).

Heart rate turbulence was primarily applied to risk stratification after myocardial infarction and to the prediction of risk of heart failure. A list of other applications with future directions is provided in a recent review [127].

Poincaré plot features

These features are based on a fitted ellipse to the Poincaré plot, which displays sample **n** of a time series as a function of sample **n-1**. These features can be seen as measures of nonlinear autocorrelation. If successive values in the time series are not linearly correlated, there will be a deviation from a line that is often properly modeled using an ellipse. The different features involve the centroid of the ellipse, the length of the two axes of the ellipse, the standard deviation in the direction of the identity line, the standard deviation in the direction orthogonal to the identity line and their combination. Those features have been successfully applied in the assessment of cardiac contractility [125], the classification of the severity of Parkinson's disease [67], the tracking of aging [62] and the assessment of chronic heart failure [102].

Recurrence plot features

A Recurrence plot is a visual representation of all the possible distances between the points constituting the phase space of a time series. There are four main elements characterizing a recurrence plot: isolated points (reflecting stochasticity in the signal), diagonal lines (index of determinism) and horizontal/vertical lines (reflecting local stationarity in the signal). The combination of these elements creates large-scale and small-scale patterns from which is possible to compute several features, mainly based on the count of number of points within each element. Table 4 provides a list of the main features used in clinical applications; for further features refer to Marwan et al. [96].

Table 4 - List of Recurrence plot features

Short name	Description
%Recurrence	Represents the number of recurrences among the total number of possible recurrences.
%Determinism	Represents the number of recurrence points that are part of a diagonal of at least $\mathbf{l_{min}}$ elements (usually $\mathbf{l_{min}=2}$), divided by the number of recurrences.
%Laminarity	Corresponds to the computation of %Determinism, but considering vertical (or horizontal) lines instead of diagonal lines.
Trapping Time	Represents the number of recurrence points that are part of vertical (or horizontal) lines, divided by the number of vertical (or horizontal) lines.
%Determinism/ %Recurrence	This ratio is often used to detect transitions of the dynamics.
Segment lengths	The mean and maximum lengths of diagonal and vertical (or horizontal) lines.
Shannon entropy	Same feature presented in the Informational domain, but applied to the distributions of the lines (either diagonal or vertical/horizontal).
Kolmogorov entropy	The slope of the line fitting the log-log plot of the distribution of diagonal lines as a function of the threshold used to create the recurrence plot.

Recurrence plots have found applications in the prediction of cardiac arrhythmia [75], event detection in EEG [96], diabetic autonomic dysfunction assessment [96,140], evaluation of the postural instability in patients affected by Parkinson [141], and are applicable to the assessment of many other physiological conditions [142].

Spatial filling index

Based on the transformation of the time series in the phase space, this feature is extracted by first dividing this space into a set of identical hypercubes. The Spatial filling index represents a measure of the number of points that on average are inside a hypercube, and describe how much the attractor fills the phase space. SFI was applied for the classification of normal heart rhythm, arrhythmia, supraventricular arrhythmia and congestive heart failure [108].

Energetic

The Energetic domain describes the features related to the energy or the power of the data. For a summary of the key concepts refer to Table 7.

Energy operators

These features combine the information of multiple points of the time series with nonlinear operations. Two features of this type are the *Plotkin and Swamy operator*, and *Teager's operator*. The latter is a specific case of the former; it is more sensitive to noise but mathematically linked to the energy of a sine wave, defined as a function of the product of its amplitude and its frequency. This fact motivates the name of these features. Despite the fact that these features were developed for sinusoidal signals, in clinical settings their application has been extended to signals showing periodicities: the Plotkin and Swamy operator was applied to segmentation and feature extraction of EEG signals [122], and the Teager's operator to the detection of peaks for the estimation of foetal heart rate from phonocardiographic signals [123].

Frequency and Time-Frequency features

The Frequency features are the ones that can be extracted from any Frequency transformation. The standard approach is to divide the spectrum in different bands (frequency ranges), and calculate the central frequency in each range, i.e. the peak in the power spectrum [22]. Beside the maximum value, other statistics include the integral over specific bands as well as different types of ratios between the cited features. For example, in the case of Heart rate variability the power spectrum is usually divided into four bands with the following frequency ranges: Ultra low frequencies < 0.003 Hz, Very low frequency 0.003-0.04 Hz, Low frequencies 0.04-0.15 Hz, and High frequencies 0.15-0.4Hz. Common features involve the power in the low and high frequency ranges, and their ratio [22,25]. The Time-Frequency transformation is nothing else than the evolution of Frequency transformations for nonstationary signals. Therefore, the Time-Frequency features are analogous to the Frequency features.

Those types of features are among the most used, and their applications involve among others the assessment of the risk of myocardial infarction and diabetic neuropathies [22,25], the characterization of heart rate variability during haemodyalisis [79], the assessment of cardiac contractility [125], the classification of cardiac death [82] and the severity of Parkinson's disease [67].

Multiscale time irreversibility

Multiscale time irreversibility quantifies the degree of temporal asymmetry of a signal, i.e. how much energy is dissipated during the development of the process. In time series analysis, this corresponds to the modification of the statistical properties of a signal under the operation of time reversal. Multiscale time irreversibility is based on the computation

of several differential time series derived from the original one. That means taking a sample at time **i+j** and subtracting from it the value of the sample at time **i**, then repeating this for all the samples **i**; the process has to be repeated with different values of **j** (representing the scale parameter), creating a set of time series. The percentage of difference between the increments and decrements inside each time series is computed; the sum of all these percentages gives the Multiscale time irreversibility [136]. Even if created for the study of heart rate variability, this measure has not been fully characterized in specific clinical settings.

Informational

The Informational domain describes the degree of irregularity/complexity of a system. For a summary of the key concepts, refer to Table 8.

Approximate entropy, Sample entropy and Fuzzy entropy

Approximate entropy (AP) is the negative natural logarithm of the conditional probability that a dataset of length N, having repeated itself for **m** samples within a tolerance **r**, will repeat itself again for one extra sample. A window of length **m** is run along the signal to generate a set of data vectors of length **m**. One then computes the number of times that the Euclidian distance between all pairs of these vectors is less than a threshold **r**. This is repeated for windows of length **m** + 1, and the logarithm of the ratio of these two numbers is taken. AP is the precursor of Sample entropy (SE); the difference is that SE excludes the counts where a vector is compared with itself. This avoids the bias these self-matches introduce in the estimation. Given this improvement, SE should be always preferred to AP. To avoid the sensitivity to the threshold **r**, a new entropy called Fuzzy entropy (FE) was developed. All the computational steps are the same, with the

difference that SE uses **r** to produce a binary classification of the distance between two windows (zero if they are more distant than **r**, one otherwise), while FE uses a fuzzy membership function to evaluate the distance. This continuous function scores as 1 if the distance between two windows is infinitesimal and decays exponentially to zero the more distant are the vectors. This improvement avoids the discontinuity of the binary classification, therefore lowering the sensitivity to the threshold. Chen et al. analyzed the three entropies on synthetic datasets, demonstrating the improved performances of FE on the characterization of synthetic datasets with different mixtures of determinism and noise [111].

AP and SE have been applied to several fields like the analysis of the EEG [109], the fluctuations of the Center of Pressure in static posturography [110], the characterization of Heart Rate variability in patients undergoing sepsis shock [62] and for the classification of multiple heart disorders [108]. On the contrary, Fuzzy entropy has been applied mainly in the characterization of EMG signals [150].

Conditional entropy

Given two random variables **X** and **Y**, the Conditional Entropy of **Y** given **X** is defined as the difference between the joint entropy of **X** and **Y** (that is the entropy of the union of **X** and **Y**), and the entropy of **X**. To provide an example, consider a time series reconstructed in a **m**-dimensional phase space. **X** can be the set of points in the **m**-dimensional space, and the union of **X** and **Y** can be the set of points given by the reconstruction of the time series in a **m+1**-dimensional phase space. The entropy used for the calculation of the conditional entropy can be any one reported in this paper. In [82] the

conditional entropy was computed using the Shannon and the Rényi entropies for the classification of cardiac death.

Compression entropy

Compression entropy is a readjustment of a compression algorithm called LZ77. After the transformation into symbols, the sequence is compressed and the entropy is computed as a function of the ratio between the length of the compressed signal and the length of the original signal. This technique was used for the classification of ventricular tachycardia [112] and chronic heart failure [74].

Kullback-Leibler permutation entropy

The Permutation entropy is based on the representation of the time series in a symbolic phase space. Considering a single point in the phase space, the idea is to substitute to each coordinate the rank among the coordinates. For example, the vector [0.3 0.2 0.7] would be substituted to the vector [2 1 3], because 0.2 is the smallest element of [0.3 0.2 0.7], therefore it is exchanged with a 1; its largest element 0.7 is exchanged with a 3; and 0.3 is the element in between, exchanged with a 2. The process is repeated for each point in the phase space, creating a set of words. Then, the Shannon entropy of the set of words is computed, creating the so-called Permutation Entropy. A normalization using the total number of possible words achievable in that phase space representation, leads to the Kullback-Leibler permutation entropy. This technique has been applied to the characterization of foetal heart rate variability [129].

Multiscale entropy

Multiscale entropy quantifies the information content of a signal over multiple time scales [136]. Consider a non-overlapping window analysis of the original time series, where

the sample mean inside each window is computed. This set of sample means constitutes a new time series. Repeating the process **N** times with a set of window lengths starting from 1 to a certain length **N**, this will give a set of **N** time series of sample means. The Multiscale entropy is obtained by computing any entropy measure (Sample entropy is suggested) for each time series, and displaying it as a function of the number of data points N inside the window (i.e. of the scale). The authors suggested focusing on the analysis of the resulting curves rather than on extracting a single parameter. However, to assist clinical classification in the case of RR interval time series, they proposed to extract the slopes of the curve over the scales from **N** = 1 to 5, and from **N** = 6 to 20 [135]. Multiscale entropy was also applied to the comparison of sleep-wake cycles of patients with heart failure, young and elderly [135].

Predictive-based features

The idea behind these features is to use the error that a model produces when trying to fit or predict the data. Models are based on hypotheses; therefore, in the case where the distance between the signal produced by the model and the real signal (i.e. the error) is high, the time series is less compatible with the underlying hypotheses of the model. The error can be evaluated through specific cost functions (e.g. the mean squared prediction error, prediction gain, sum of the residuals) and can be referred as an index of complexity [137]. Because the modeling of physiological signals is entirely another field of research, we will not address this topic further. Examples of this approach are the usage of an Autoregressive model [138], local nonlinear predictors [137] and predictors based on conditional distribution [139] for the modeling of heart rate variability.

Shannon entropy and Rényi entropy

As classical measures taken from information theory, these entropies are estimated after the transformation of the dataset into bins or a symbolic sequence. Counting the relative frequency of each bin/word, the Shannon entropy is estimated as the sum of the relative frequencies weighted by the logarithm of the inverse of the relative frequencies (i.e. when the frequency is low, the weight is high, and vice versa). A generalized version of the Shannon entropy is the Rényi entropy, which corresponds to the sum of the inverse of the relative frequencies elevated to a certain power **q**.

In the case of the bin transformation, Shannon entropy was applied in the classification of the heart rate variability of young/elderly and patients suffering from coronary artery disease [138]. For the symbolic transformation, Shannon entropy was applied to the detection of ventricular tachycardia or fibrillation [103], the characterization of autonomic cardiac modulation during arrhythmias [101] and the evaluation of heart rate and blood pressure variability in patients with dilated cardiomyopathy [100].

Similarity indexes

The Similarity indexes are features quantifying the "distance" (i.e. difference) between two time series, according to different metrics. The two time series can be either two different time series, or two segments from one time series. Usually the distance is calculated after a qualitative transformation of the dataset, and is expressed as a combination of the probability distributions of the words/bins or a combination of the entropy of the words/bins. When a windowed analysis of a single time series is pursued, the idea is to compare either the similarity between two successive windows, or the similarity between a window for a given physiological state (e.g. normal) and all the

other windows, in order to track how the system is changing (e.g. moving from physiological to pathological). Examples of similarity indexes with associated clinical applications follow: 1) A similarity index defined as the sum of product of the bins within the distributions of each segment was used in the assessment of anesthetic depth [143]; 2) A similarity index defined as the difference between the Shannon entropies of the symbolic distributions of each segment was used in the assessment of the effect of aging, atrial fibrillation and heart failure in heart rate variability [144], and the detection of arrhythmias [99]; 3) Considering two segments after a Grid transformation as two images, it is possible to define similarity indexes based on image processing. Among the several proposed, one that worked effectively in the detection of ventricular fibrillation was the count of the pixels that were visited in both the considered images [84].

Invariant

The Invariant domain describes the properties of a system that are not supposed to change over either time or space. For a summary of the key concepts refer to Table 9.

Allan and Fano scaling exponents

Those two parameters were created to characterize the fractal properties of a point process. The idea is to count the number of events in non-overlapping windows of the signal of fixed time-length **T**. The Fano factor at time **T** is the variance of the number of events divided by the mean number of events (it also called *index of dispersion*). The Allan factor at time **T** is the ratio between the second order moment of the count differences between two successive windows, and the mean number of events. The slope of the line fitting the log-log plot of the factor as a function of **T** gives, respectively, the

Fano and the Allan scaling exponents [95,106]. These parameters were applied to the characterization of supraventricular arrhythmia and congestive heart failure [107].

Correlation dimension

The correlation sum of a time series is defined as the number of points in the phase space that are closer than a certain threshold **r** [92]. An array of correlation sum values is then created by repeating the process for a certain range of thresholds. One then computes the correlation dimension as the slope of the line fitting the log-log plot of the correlation sum as a function of the threshold. This feature gives a measure of the dimensionality of the attractor, and was used to study blood pressure and mean cerebral blood flow velocity in diabetic autonomic neuropathy [113].

Detrended fluctuation analysis

DFA is a technique quantifying how the fluctuations of a signal scale with the number of samples of that signal. First, the cumulative version of the time series is created, where the value of the original time series at a given point is replaced by the sum of values up to and including that point. The idea is to divide this cumulative time series into all the possible non-overlapping windows of length **m**, then detrend each window and compute the standard deviation within each window (i.e. the magnitude of fluctuations). Each window is detrended by removing the regression line fitting the data. The average of the standard deviation computed in each window is stored. Repeating the process for a certain set of window lengths yields a time series that may exhibit scaling behaviour, i.e. where fluctuations scale according as a power of the window length [116]. Several features can be extracted from DFA: (1) the slope of a line fitting the log-log plot of the standard

deviation as a function of the window length; (2) the local derivative in each point of the log-log plot [114].

For the Heart rate variability an additional feature arises from the division of the log-log plot in two areas to extract two scaling exponents. Indeed, in different studies, two scaling exponents more properly fitted the DFA of the HRV than the standard slope of the log-log plot. Another particular application of DFA to Heart rate variability uses, instead of the raw data, the magnitude and sign of the derivative of an RR interval time series [115].

DFA is one of the most used techniques in variability analysis, and was applied to the evaluation of posture, exercise and aging [114], sleep stage classification [115] and the prediction of sepsis [62].

Diffusion Entropy

Assuming the time series is a diffusion-type stochastic process, this measure quantifies its scaling behaviour over time. The first step is to calculate the diffusion trajectories, which are particular types of cumulative time series extracted from the original. Then, the probability distribution at a given cumulative time is built from the values of all the diffusion trajectories at that time, and the Shannon entropy is computed at that time. Repeating the procedure for all the times, a set of Shannon entropies as a function of time is created. The slope and offset of the line fitting the normal-log plot of the Shannon entropies as a function of the cumulative time gives information regarding the scaling behaviour of the system [119]. This technique was mainly applied to the characterization of congestive heart failure [102,117,118].

Embedding scaling exponent

This technique estimates how the variance of the time series reconstructed in the phase space changes with the value of the embedding dimension. The variance is computed from the diagonal of the covariance matrix of the time series reconstructed in the phase space. The slope of the line fitting the log-log plot of the variance as a function of the embedding dimension gives the Embedding scaling exponent. ESE has been applied to gait analysis of patients with Huntington's disease [121].

Finite growth rates and Largest Lyapunov exponent

After the representation of the time series in the phase space, the neighbourhood of each point in the phase space is computed. The idea is to compute how the considered point and its neighbour separate after a certain time **t**; the separation is measured using the Euclidean distance between the points. The measure used by the Finite growth rates is the ratio between the final (i.e. after **t**) and the initial distance of the two points. The repetition of this process to all the points in the phase space gives a set of ratios, whose average corresponds to the Finite growth rates (FGR). This technique was used for the detection of ventricular tachycardia or fibrillation [103].

Consider now that the Finite growth rate is computed without the normalization by the initial distance (therefore, the Finite growth rates now represent a final distance). If all the FGR for any time **t** are computed, the slope of the line fitting the log-log plot of the FGR as a function of **t** gives the Largest Lyapunov exponent. The Largest Lyapunov exponent expresses the exponential rate of divergence of the trajectories in the system, starting with two infinitely close points in the phase space [10,92]. Theoretically, the Largest Lyapunov exponent is positive only in the case of chaotic systems, but this result is not assured

when considering noisy time series. This dynamical-invariant feature was used for gait analysis [131], and congestive heart failure and atrial fibrillation evaluation [130].

Higuchi's algorithm

The time series is divided into non-overlapping segments of **m** samples. Consider the samples in the **i**-th position inside each segment. The distance between those samples can be calculated as their absolute difference. If the normalized sum of all the distances of the samples in the **i**-th position is calculated for each position, and then averaged, the result is a measure of the length of the time series. The Higuchi's algorithm involves computing the length of the time series over a certain range **m**. The slope of the line fitting the log-log plot of the lengths as a function of 1/**m** gives a measure of how the length of the time series scales. This technique was applied to detect events in EEG [128], and characterize heart rate variability in chronic heart failure patients [102].

Index of variability

The Index of variability is based on a Point process model. The time series is divided in non-overlapping windows of a given length, and for each window the number of events is counted. Then, the variance of the number of events for that window length is computed. The procedure is repeated for different window lengths. The scaling behaviour is extracted from the derivative of the variance of the number of events as a function of time. Index of variability was applied to evaluate the scaling properties of traffic processes in network engineering [61].

Kolmogorov-Sinai entropy

The Kolmogorov-Sinai entropy is a measure expressing the information needed to predict which part of the phase space the dynamics will visit at a time **t+1**, given the trajectories

up to time **t** [92]. Consider the natural logarithm of the ratio of the Correlation sum computed for an embedding dimension **m**, with the Correlation sum computed for an embedding dimension **m+1**. The plateau of the curve of this ratio as a function of the logarithm of the threshold gives the KS entropy. This measure was used to study blood pressure and mean cerebral blood flow velocity in diabetic autonomic neuropathy [113].

Multifractal exponents

Multifractal time series are characterized by a set of different local scaling exponents (i.e. different scaling exponents at different times). One of the most used techniques to evaluate multifractality is the Wavelet transform [133]. The idea is to generate the Wavelet transform by using a particular type of mother Wavelet, called Wavelet leader, and from it extract the multifractal spectrum. This method was applied to the characterization of heart failure [102,133], and intrapartum diagnosis of fetal asphyxia [134].

Power spectrum scale exponent

This feature represents a classical way to compute the scaling exponent of a time series, and is computed as the slope of the line fitting the log-log plot of the power spectrum of the time series as a function of the frequency. This gives a reliable estimate of the scaling exponent of the time series only when the underlying process belongs to the fractional Gaussian noise family, and only for a specific range of the so-called Hurst exponents; otherwise, the use of this feature should be avoided [116].

Scaled windowed variance

Scaled windowed variance involves the same mathematical steps as DFA, with the difference that the original times series is studied, rather than its cumulatively integrated

version. Despite the similarity, Scaled windowed variance is preferable to DFA when the underlying process belongs to the fractional Brownian motion family [116].

Rescaled detrended range analysis

Rescaled detrended range analysis is similar to Detrended fluctuation analysis. Instead of calculating the standard deviation on each window, Rescaled detrended range analysis computes the "range", defined as the difference between the maximum and the minimum values of the samples in the window. RDRA is preferable to DFA when the underlying process belongs to the family of persistent fractional Gaussian noise [116].

Challenges

The wide array of techniques discussed in our review represents a starting point for advanced variability analyses. Clearly, more investigations are required to untangle all the information hidden within biological signal variability. Two grand challenges stand out, which we believe represent important future goals of variability analysis:

1) Reduction of the dimensionality of variability analysis for a practical use in clinical settings. Biological variability and variable physiological conditions motivate the use of different descriptors, capable of describing in the most accurate way possible the system under study. However, the collection of this huge array of univariate and multivariate techniques calls for the need to personalize the relevant information for each single subject, considering specific stress conditions that could be either pathological or not (e.g physical exercise). A step toward this result would be given by the characterization of the nature and degree of interdependence between measures of variability in specific clinical situations. Few studies have tried to face this challenge in a broad way. An example, such a characterization has been provided by Cerutti et al (2007) to address this problem;

however, they focused only on the analysis of a small set of nonlinear, univariate heart variability techniques applied to the assessment of heart failure [102]. We believe that a more extensive characterization would enable the selection of the optimal set of descriptors for a given clinical situation, thereby reducing the complexity of the studies while still preserving all the valuable information. This would form the basis for an advanced type of data reduction and fusion technique, with the final aim to converge the different pieces of valuable information in an easy-to-use tool for physicians and nurses.

2) Extension of variability analysis to multivariate techniques characterizing multi-organ variability and connectivity. In the scientific community it is well accepted that the human body is a complex system, and its behavior depends on the interactions of its components [15,151–153]. In spite of this fact, the common approach is to use univariate techniques [154,155], and this review is a proof of it. Examples of multivariate studies exist, however they are mainly focused on the evaluation of interactions between at maximum two systems, like the cardiac and respiratory systems through frequency analysis (spectral, Wavelet…) [156–158], or the evaluation of the differences between heart rate variability and blood pressure [159] or pulse oximeter data [79]. Multiorgan variability and connectivity, respectively defined as the ensemble of patterns of variation over time and interconnection over space, represent new tools to describe the behavior of the human body as a whole[17]. These tools will likely provide complementary information to that given by standard univariate techniques, and the simultaneous use of the two approaches together holds the promise of new clinically valuable results.

Conclusion

Physiological time series are complex entities characterized by a variety of properties, some of which are often intertwined, such as through a combination of stochasticity and determinism. This complexity demands the use of a vast array of techniques, capable of describing in the most accurate way possible the time series under study. In this article more than 70 variability techniques were presented, explaining with an intuitive approach the idea behind them, discussing their limitations and associating references related to their clinical applications. Moreover, a novel way of classifying these variability techniques was proposed, specifically according to whether they perform transformations of the data or extract features from the data (Table 2). The paper further highlights two grand challenges that the scientific community will face in the context of variability analysis for clinical applications.

Competing Interests

Andrew J.E. Seely founded Therapeutic Monitoring Systems in order to commercialize patented Continuous Individualized Multiorgan Variability Analysis (CIMVA) technology, with the objective of delivering variability-directed clinical decision support to improve quality and efficiency of care.

Author's Contribution

AB analysed current literature and prepared the manuscript. AB proposed the classification of the variability techniques, with the supervision of AJES and AL. AJES

and AL supervised, revised and gave the final approval of the manuscript. All authors read and approved the final manuscript.

Table 5 - Summary of the Statistical domain

Domain Assumptions	Features	Feature Assumptions	Transformation used	References
The information is held in the statistical descriptors of the data distribution (stochastic process)	Form factor	Gaussian distribution of the time series and its derivatives		124
	Lee parameter	Gaussian distribution of the time series		132
	Symbolic dynamics features		Symbolic dynamics	82,99–103
	Standard statistical features	Gaussian distribution of the time series	Bins	22,25,67,74,82,125
	Turns count	Spike-like behaviour of the time series		124

Table 6 - Summary of the Geometric domain

Domain Assumptions	Features	Feature Assumptions	Transformation used	References
The information is held in the specific position of the points represented in a reference space (deterministic system)	Grid counting	Low dimensionality of the attractor	Grid transformation	84,126
	Dynamical moments	Deterministic trends can be effectively described through statistical methods	Phase space representation	120
	Heart rate turbulence	Makes use only of ectopic beats		127
	Poincaré plots features	Low dimensionality of the attractor	Poincaré plots	62,67,102,125
	Recurrence plot features	The attractor can be effectively described by the distances between its points	Recurrence plots	75,96,140–142
	Spatial filling index	The attractor is described by its degree of sparseness	Phase space representation	145

Table 7 - Summary of the Energetic domain

Domain Assumptions	Features	Feature Assumptions	Transformation used	References
The information is held in the energy of the signal	Frequency features	Stationarity of the time series	Frequency transformations	22,25,67,79,82,125
	Time-Frequency features		Time-frequency transformations	22,25,67,79,82,125
	Energy operators	The time series expresses periodicity		122,123
	Multiscale time irreversibility	The time series is created from a dissipative process, which dissipation can be quantified by the degree of temporal asymmetry		136

Table 8 - Summary of the Informational domain

Domain Assumptions	Features	Feature Assumptions	Transformation used	References
The information is held in the degree of	Approximate entropy			62,108–110
	Conditional entropy		Bins	82

complexity, therefore: distance from periodicity and stochasticity, distance from a reference model, distance from a precedent pattern within the data	Compression entropy			[74,112]
	Fuzzy entropy			[111]
	Kullback-Leibler permutation entropy		Symbolic dynamics and Phase space representation	[129]
	Multiscale entropy	The complexity changes depending on the window length used in the analysis		[135]
	Predictive-based features	The data follows a model, and the deviation (prediction error) from that model describes changes in the system.	Multiple	[137–139]
	Sample entropy			[62,108–110]
	Shannon entropy		Bins	[100,101,103,138]
	Similarity indexes	The comparison of the properties of two successive windows allows the detection of changes in a time series	Multiple	[84,99,143,144]
	Rényi entropy		Bins	[82]

Table 9 - Summary of the Invariant domain

Domain Assumptions	Features	Feature Assumptions	Transformation used	References
The information is held in those properties of the time series that are invariant – i.e. not supposed to change over either time or space.	Allan scaling exponent	The time series is modeled as a point process, and the ratio between the second order moment of the difference between the number of events of two successive windows and the mean number of events has scale-invariant properties	Point process	[95,106,107]
	Detrended fluctuation analysis	The standard deviation of the detrended cumulative time series has scale-invariant properties		[62,114,115]
	Diffusion Entropy	The time series is modeled as a family of diffusion processes, which Shannon entropies has scale-invariant properties		[102,117,118]
	Embedding scaling exponent	The variance of the attractor at different embedding dimensions has scale-invariant properties	Phase space representation	[121]
	Fano scaling exponent	The time series is modeled as a point process, and the variance of the number of events divided by the mean number of events has scale-invariant properties	Point process	[95,106,107]
	Higuchi's algorithm	The length of the time series at different windows has scale-invariant properties		[102,128]
	Index of variability	The time series is modeled as a point process, and the variance of the number of events has scale-invariant	Point process	[61]

	properties			
Multifractal exponents	Multiple scaling exponent characterize the time series	Wavelet transform		[102,133,134]
Power spectrum scaling exponent	Stationarity, the power spectrum follows a $1/f^b$ like behaviour	Power spectrum		[116]
Rescaled detrended range analysis	The range (difference between maximum and minimum value) of a time series has scale-invariant properties			[116]
Scaled windowed variance	The standard deviation of the detrended time series has scale-invariant properties			[116]
Correlation dimension	The time series is extracted from a dynamical system, and the number of points in the phase space that are closer than a certain threshold has scale-invariant properties	Phase representation	space	[113]
Finite growth rates	The time series is extracted from a dynamical system, which is described by its dependence on the initial conditions (how two points that are close in space and time separate after a certain amount of time) – the ratio between the final and the initial time is an invariant of the system	Phase representation	space	[103]
Kolmogorov-Sinai entropy	The time series is extracted from a dynamical system, and it is possible to predict which part of the phase space the dynamics will visit at a time t+1, given the trajectories up to time t	Phase representation	space	[113]
Largest Lyapunov exponent	The time series is extracted from a dynamical system, which is described by its dependence on the initial conditions (how two points that are close in space and time separate after a certain amount of time) – the distance grows on average exponentially in time, and the exponent is an invariant of the system	Phase representation	space	[130,131]

Chapter 3

3 Dimensionality of variability

3.1 Commentary

The extraction of multiple measures of variability from a physiological signal is associated with the technical challenge of dealing with the large dimensionality (i.e. number of variables) of the dataset. This is true both when the data is used for statistical test of hypothesis and to create a mathematical model.

In statistical testing the objective is the evaluation of a given null hypothesis, e.g. equality of the median value of a given measure of variability in two populations (say healthy and diseased). This evaluation can be applied to all the measures of variability under study, thereby producing what in statistical terms are called multiple comparisons. Given that in this example the number of comparisons is identical to the number of measure of variability, there is the problem that rejections of the null hypothesis could occur simply due to chance. Statisticians developed a variety of techniques for multiple comparison correction, such as Bonferroni, Holm-Bonferroni, and false discovery rates[160,161], which goal is the control of type I errors (i.e. false rejections of the null hypothesis). This however comes at the price of lack of control for type II errors (i.e. false acceptance of the null hypothesis), with the effect that clinically relevant variables may be failing to surpass the threshold for significance. This is the same trade-off usually seen in the comparison of diagnostic tests with screening tests. A diagnostic test is used to identify the presence of a disease or condition, and therefore needs to have lower type I errors (i.e.

false positives). Conversely, a screening test is used to guide whether a diagnostic test should be offered to patients at high risk of having a disease or condition, and therefore needs to have lower type II errors (i.e. false negatives).

In mathematical modeling the relationship between the variables is commonly used for one of three tasks: 1) clustering (i.e. identification of subgroups within the data showing similar properties), 2) classification (i.e. attribution the data into a set of pre-defined categories), regression (i.e. mapping the data into a continuous variable). For each task there is a large set of different models available (e.g. neural networks, support vector machines, decision trees, etc..[162,163]). Regardless of its type, the model will be dependent on a set of parameters whose estimation requires the solution of an optimization procedure. When the number of variables used in the model is too large respect to the number of elements in the dataset, the optimization procedure will either not converge or converge to a model that will fit the given dataset, yet will not be able to be valid for a new dataset (i.e. will lack generalization capabilities).

Therefore, no matter the type of analysis done, whenever possible it is ideal to reduce the dimensionality of the data through data reduction techniques[164]. The following article highlights the required steps to transform multiple physiological signals into a single score.

3.2 Manuscript 2: System and method for generating composite measures of variability

This manuscript is a patent published in WIPO Patentscope in date 10 January 2013. A peculiarity of patents is that the elements of a figure are referenced with numbers which are then used in the body of the manuscript.

My contribution: creation of the processing scheme for the creation of composite measures, writing of the patent, review and final approval.

SYSTEM AND METHOD FOR GENERATING COMPOSITE MEASURES OF VARIABILITY

Andrew JE Seely, Andrea Bravi, André Longtin

ABSTRACT

A system and method are provided for assessing the degree of critical illness of patients in Intensive Care Units (ICUs), and the detection of occurrence of specific clinical outcomes, pursued through the monitoring of a variety of physiological signals, extraction of multiple measures of variability, and integration of the relevant information the measures of variability bring. The integration process leads to the creation of a compound variable, referred to as a composite measure of variability, which can be tuned to various clinical needs, since the process of normalization, feature selection and data integration are customizable to the clinical indication, planned use of the composite variability monitoring. The system and method also enable joining the composite measure of variability with clinical factors, so as to estimate the likelihood of developing specific clinical outcomes.

TECHNICAL FIELD

The following relates generally to generating composite measures of variability, and more particularly to generating composite measures of variability for clinical applications.

[0001] BACKGROUND

There is an increasing interest in the application of variability monitoring to improve clinical outcomes. For example, there has been extensive research into the development of multiple measures of variability and the characterization of how variability changes with respect to specific stimuli. However, several steps need to be made to create a monitoring system capable of tracking the physiological conditions of patients undergoing different types of stresses, either physiological (e.g. physical exercise) or pathological. In particular, the following challenges need to be addressed.

The first challenge is in the characterization of the nature and degree of interdependence between measures of variability in specific clinical situations. The common approach in the literature is reductionist, wherein just a few variability measures, often of the linear type, are selected among the huge array that the literature currently offers. Heart rate variability is a good example of a phenomenon that has been studied with a few nonlinear techniques at a time, such as the Recurrence Plot and Poincare plot, which reveal phase space structure; Point-process techniques, which focus on the stochastic nature of the underlying process; and Multi-Scale Entropy, which investigates regularities across different time scales. However, it is a well-accepted idea that there are many domains of variability (i.e. time, frequency, scale-invariant, etc…) holding clinically valuable information, and that a better assessment of clinical situations requires joining the many

pieces of information that each single technique can provide. A broad comparison of the variability techniques accessible today in specific clinical settings would enable an understanding of the independent value of each and of their overlap. It would also simplify the decision of choosing a particular set of techniques, thereby providing the basis for an adaptable system for variability monitoring.

The second challenge is in the extension of variability analyses to multivariate techniques characterizing multi-organ variability. The common approach is to use linear uni-variate techniques, often combined with a small set of nonlinear uni-variate techniques. Examples of multivariate studies exist, and are mainly focused on the evaluation of interactions between the cardiac and respiratory systems through frequency analysis (spectral, wavelet…), or the evaluation of the differences between heart rate variability and blood pressure, or pulse oximeter data, which measures the oxygen saturation in the blood. A challenge that variability research needs to face is to consider the different components of a physiological system not independently, but as a whole. Several multivariate techniques are accessible but their application remains rare in clinical settings. Examples are the multivariate state space representation, Cross Recurrence Plots, Cross Fuzzy Entropy, and Multidimensional Modeling.

It is an object of the following to address the above-noted challenges.

SUMMARY

In one aspect, there is provided a method comprising: obtaining a plurality of measures of variability; and combining the plurality of measures of variability to obtain a composite measure of variability.

In another aspect, there may be provided a computer readable medium comprising computer executable instructions for performing the above method.

In yet another aspect, there is provided system comprising a processor coupled to a memory, the method comprising computer executable instructions for performing the above method.

BRIEF DESCRIPTION OF THE DRAWINGS

Embodiments will now be described by way of example only with reference to the appended drawings wherein:

Figure 5 is a block diagram of a system for generating composite variability measures.

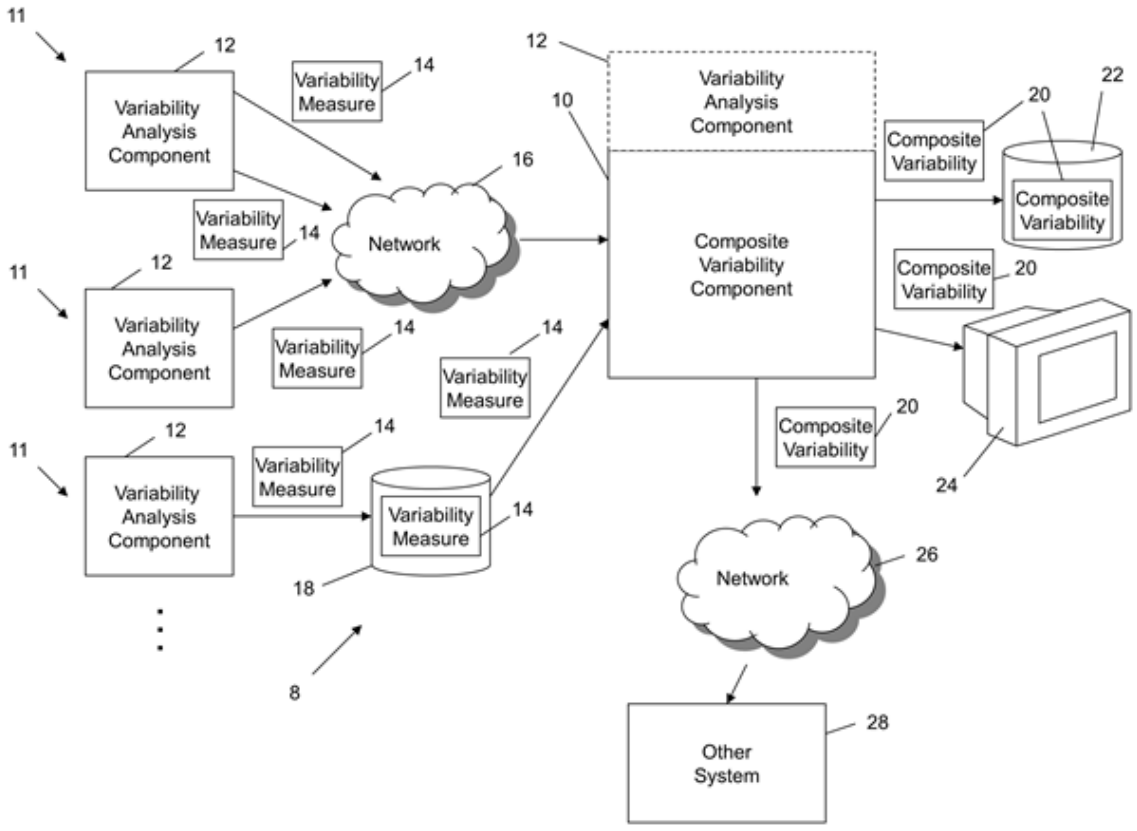


Figure 5 - Block diagram of a system for generating composite variability measures

Figure 6 is a block diagram of a hospital site including a variability analysis server.

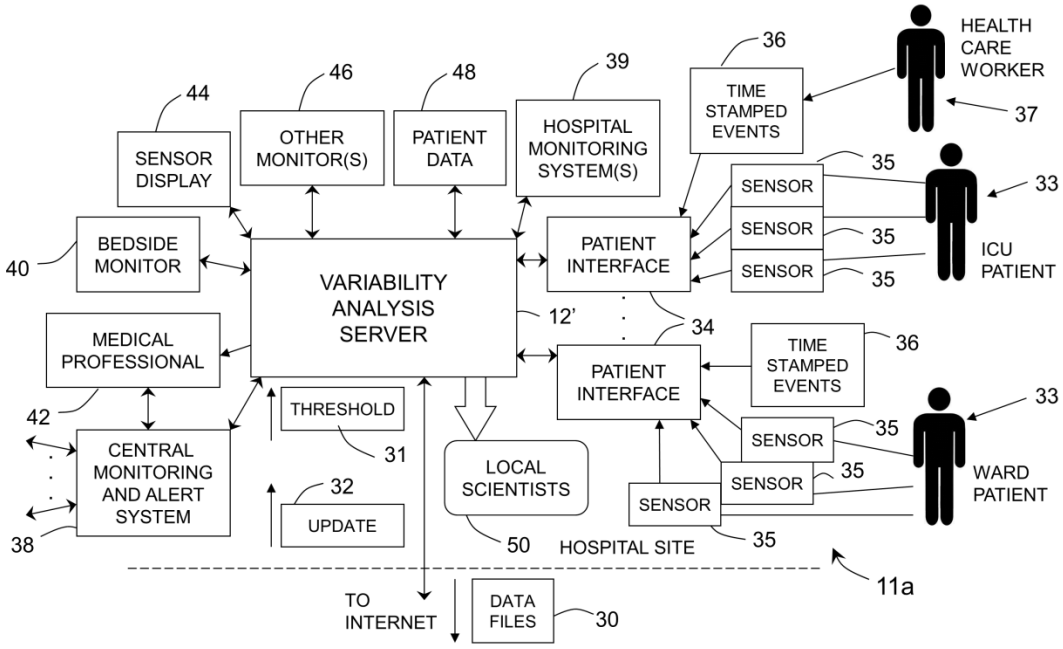


Figure 6 - Block diagram of a hospital site including a variability analysis server

Figure 7 is a block diagram of a clinic site including a variability analysis server.

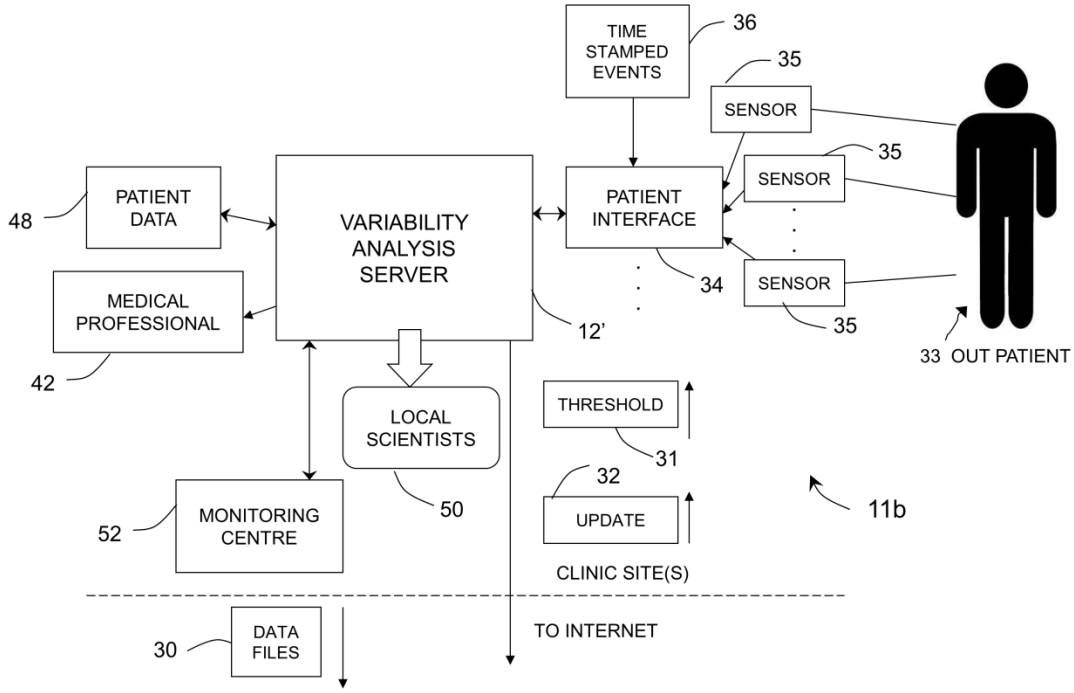


Figure 7 - Block diagram of a clinic site including a variability analysis server

Figure 8 is a block diagram of a mobile site including a variability analysis server.

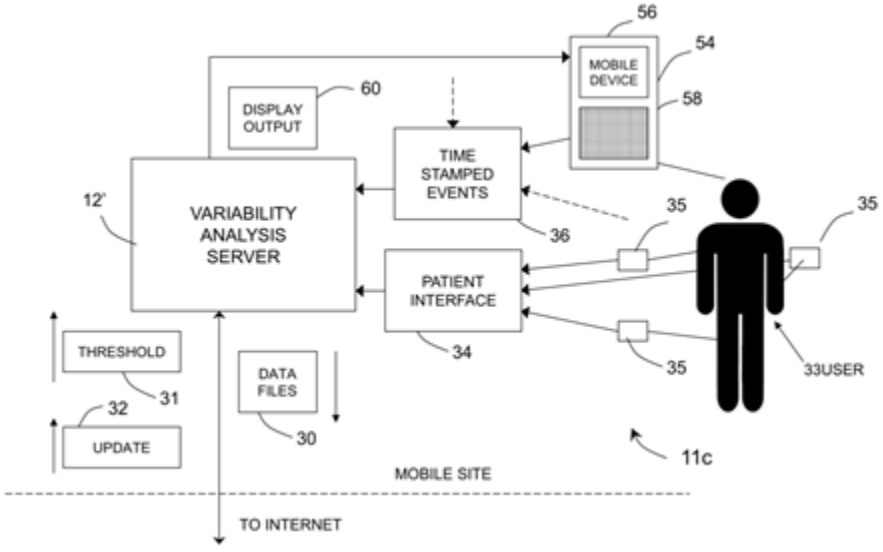


Figure 8 - block diagram of a mobile site including a variability analysis server

Figure 9 is a block diagram illustrating an example of a process flow for generating a composite variability output.

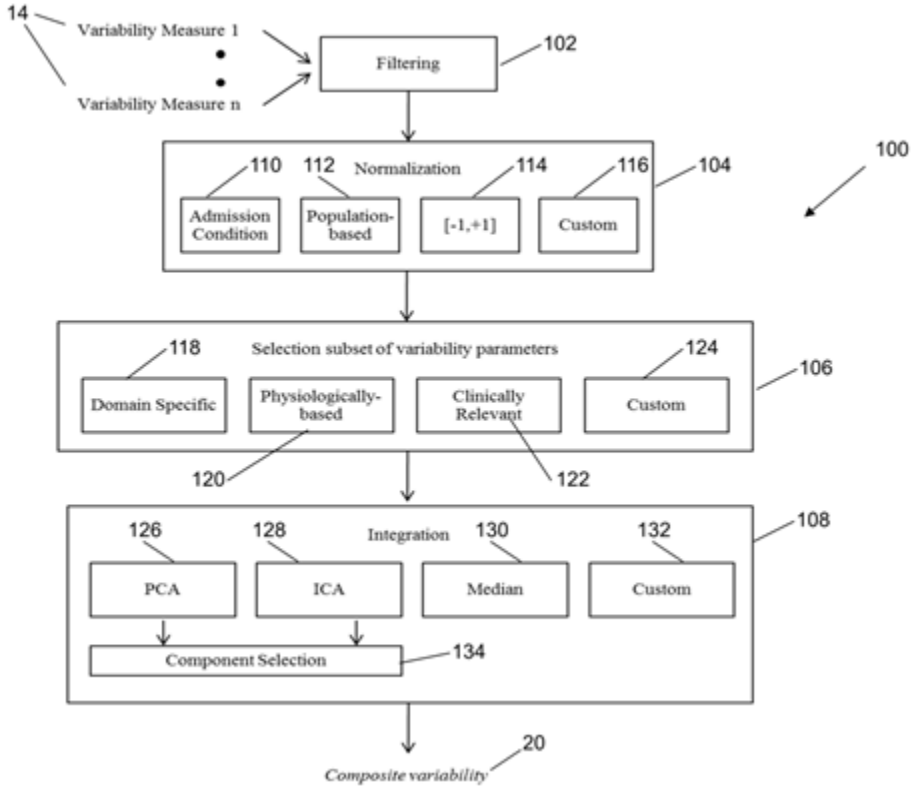


Figure 9 - Process flow for generating a composite variability output

Figure 10 illustrates a set of variability time series graphs.

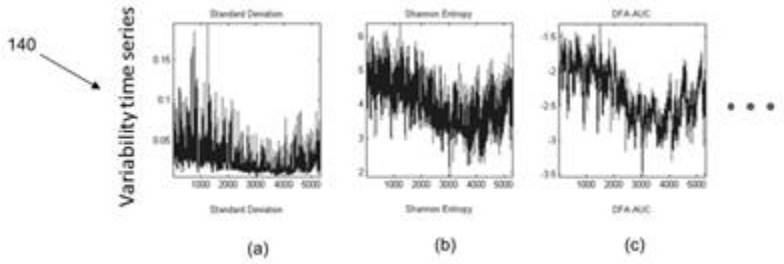


Figure 10 - Example of variability time series graphs

Figure 11 illustrates a set of filtered variability time series graphs.

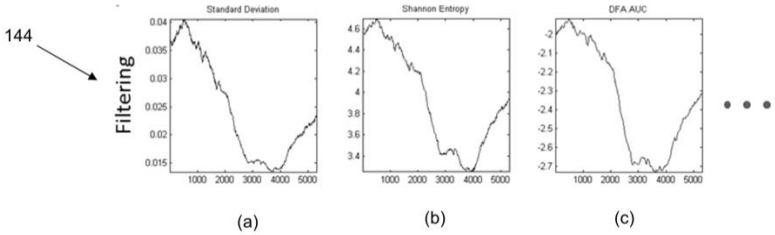


Figure 11 - A set of filtered variability time series graphs

Figure 12 illustrates a set of normalized variability graphs.

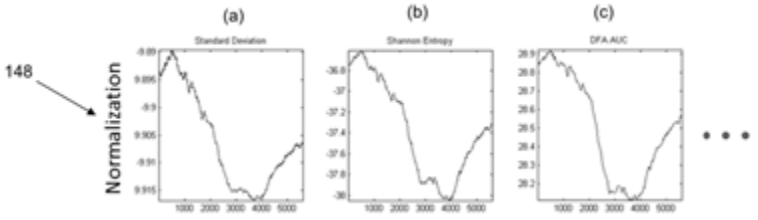


Figure 12 - A set of normalized variability graphs

Figure 13 illustrates an integration of the normalized variability graphs of Figure 12.

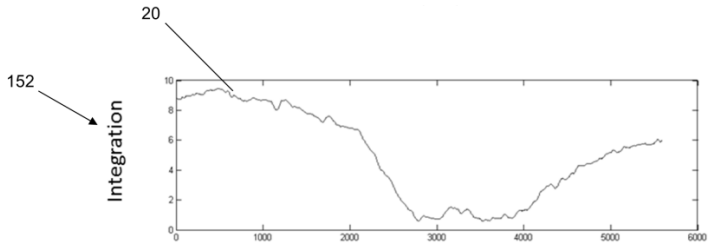


Figure 13 - Integration of normalized variability graphs

Figure 14 is two graphs illustrating a population-based normalization.

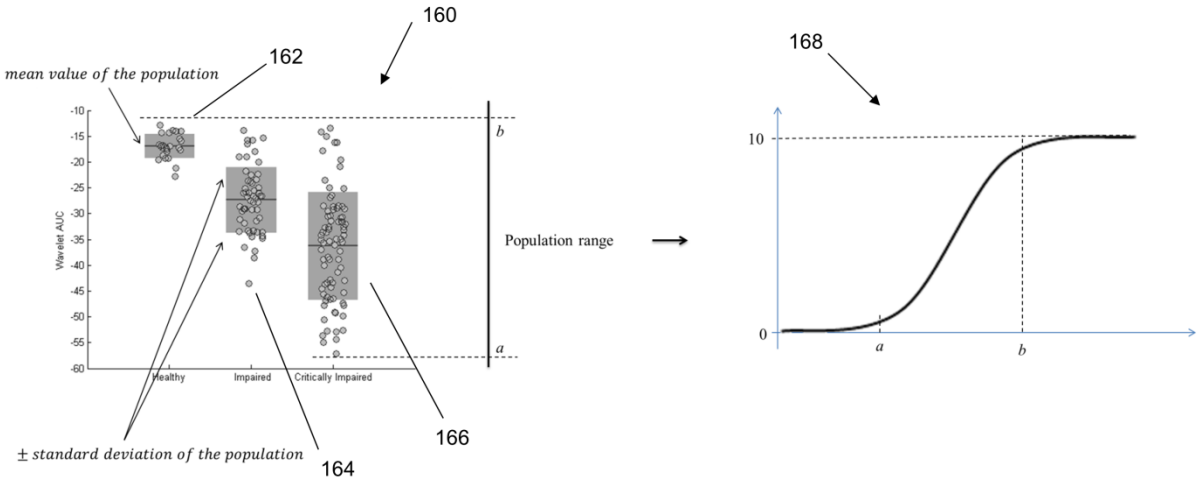


Figure 14 – Population-based normalization

Figure 15 is a flow chart illustrating an example of a set of computer executable operations that may be performed in using a composite variability output to determine a clinical likelihood.

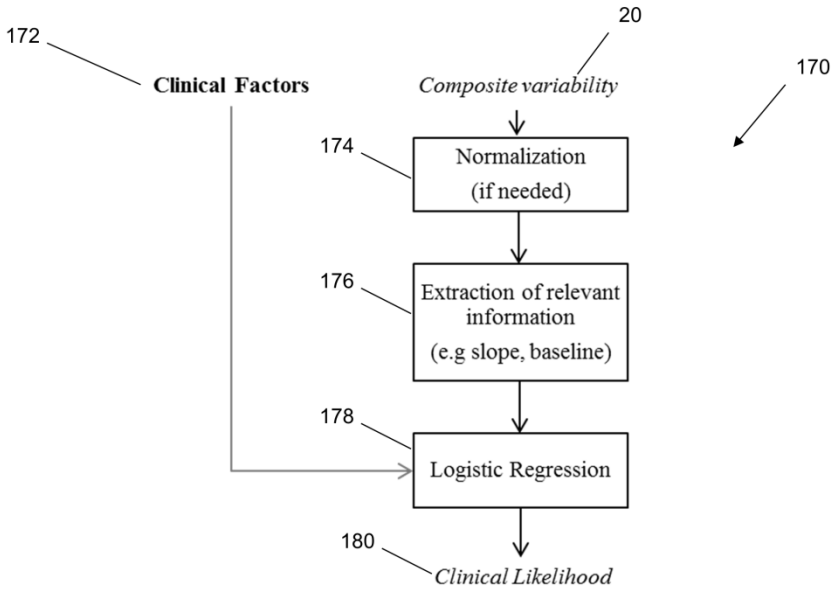
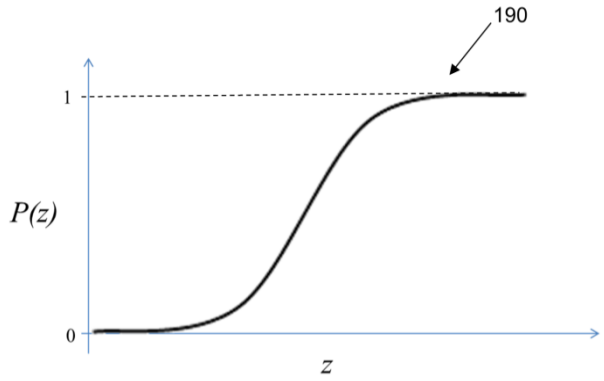


Figure 15 – From variability to clinical likelihood

Figure 16 illustrates a generic logistic regression model.



$$P(z) = \frac{1}{1 + e^{-z}}$$

$$z = \alpha + \sum \beta_j X_j$$

Figure 16 - Logistic regression model

Figure 17 is a chart illustrating an example of a Principal Component Analysis (PCA) composite variability trend output tracking the development of sepsis wherein time t=0 corresponds to the diagnosis of sepsis.

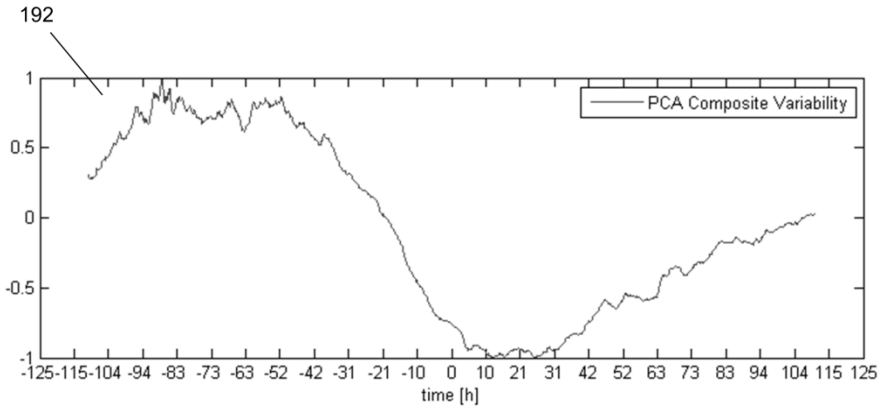


Figure 17 - Example of a Principal Component Analysis

Figure 18 is a chart illustrating an example of a median composite variability trend output during the monitoring of a patient who failed extubation, pre- and during a spontaneous breath trial wherein time t=0 corresponds to the beginning of the spontaneous breath trial.

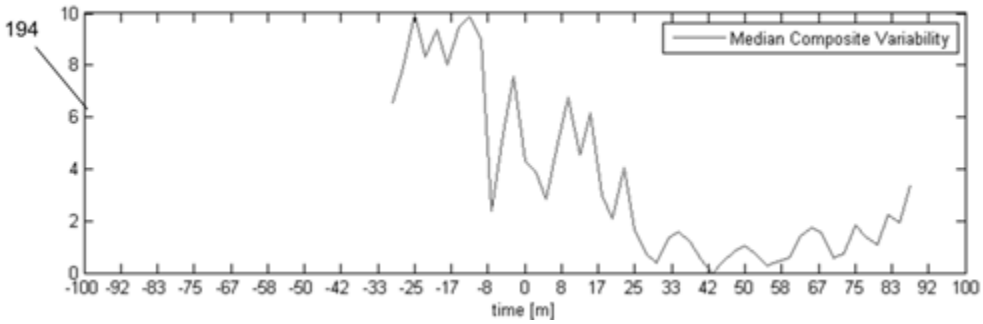


Figure 18 - Example of a median composite variability for a failed extubation

Figure 19 is a chart illustrating an example of a median composite variability trend output during the monitoring of a patient who succeeded extubation, pre- and during a spontaneous breath trial wherein time t=0 corresponds to the beginning of the spontaneous breath trial.

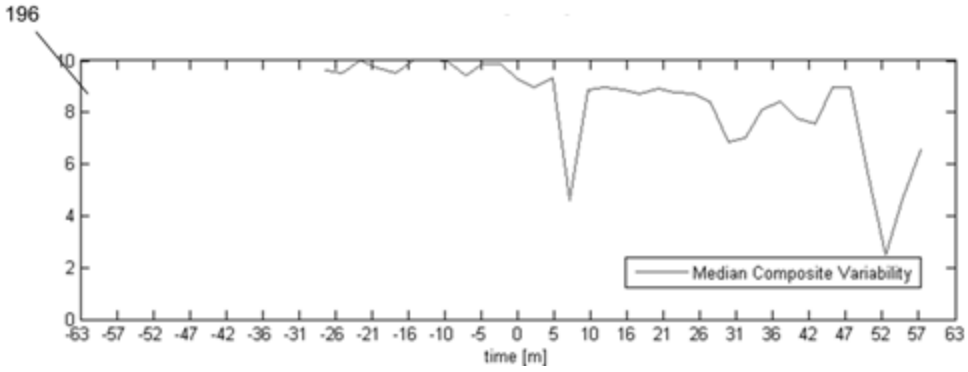


Figure 19 - Example of a median composite variability for a passed extubation

Figure 20 is three examples of population trends: Shannon Entropy (shannEn) measure of variability (left); Plotkin-Swamy engery operator (PSeo) measure of variability (center); Fuzzy Entropy (fuzEn) measure of variability (right)

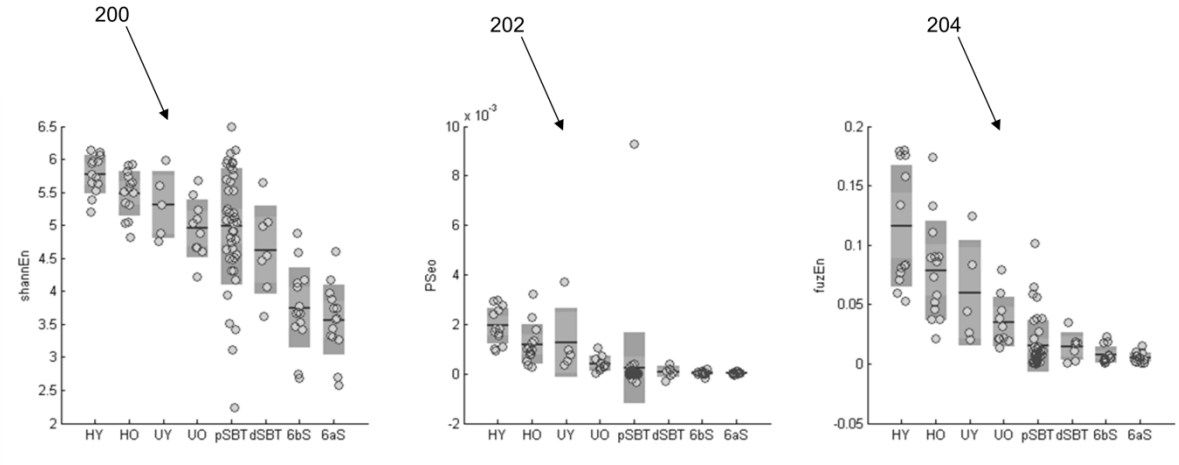


Figure 20 - Population trends

DETAILED DESCRIPTION OF THE DRAWINGS

It will be appreciated that for simplicity and clarity of illustration, where considered appropriate, reference numerals may be repeated among the figures to indicate corresponding or analogous elements. In addition, numerous specific details are set forth in order to provide a thorough understanding of the example embodiments described herein. However, it will be understood by those of ordinary skill in the art that the example embodiments described herein may be practised without these specific details. In other instances, well-known methods, procedures and components have not been described in detail so as not to obscure the example embodiments described herein. Also, the description is not to be considered as limiting the scope of the example embodiments described herein.

Composite Measures of Variability for Clinical Applications

In addition to the above described challenges, it has been recognized that a third challenge needs to be addressed in tracking the physiological conditions of patients undergoing different types of stresses. The third challenge is in the creation of a composite measure of variability to detect critical changes in health status. The collection of several independent measures of organ and multi-organ variability calls for easy-to-use metrics describing the physiological condition of a patient. Considering clinical settings, the application of this type of analysis represents an old problem in fields like EEG analysis, but a new one in the case of multi-organ variability analysis.

The following relates to the assessment of the degree of critical illness of patients in Intensive Care Units (ICUs), and the detection of occurrence of specific clinical outcomes, pursued through the monitoring of a variety of physiological signals,

extraction of multiple measures of variability, and integration of the relevant information the measures of variability bring. As will be discussed below, the integration process leads to the creation of a compound variable, which can be tuned to various clinical needs, since the process of normalization, feature selection and data integration are customizable to the clinical indication, planned use of the composite variability monitoring. The compound variable may be hereinafter referred to as a composite measure of variability. The following also presents how to join the composite measure of variability with clinical factors, so as to estimate the likelihood of developing specific clinical outcomes.

An object of the following is therefore to provide a procedure unifying the clinically relevant information from multiple measures of variability.

It can be appreciated that the measures of variability that are applicable to the following techniques include both single-organ and multi-organ variability measures. Furthermore, the following provides a procedure taking into account the information extracted from a single physiological signal, as well as a variety of physiological signals. These include, without limitation: electrocardiogram, end tidal capnography, oxygen saturation and arterial blood pressure waveforms, and their derived time series (e.g. R-R interval, inter-breath interval, pulse transit time). As such, the following joins clinically relevant information from multiple variability signals extracted from a variety of physiological signals, creating a composite measure of variability.

The composite measure of variability may be created by applying the following procedures: filtering, normalization, selection of relevant information, and integration. The principles described herein provide for the development of a composite measure of

variability which is tunable depending on the clinical needs. In particular, the types of normalization, selection, and integration can be chosen according to a purpose of use.

The proposed filtering methods may include, without limitation, Savitzky-Golay filter, Butterworth filter, Bessel filter, elliptic filter and Chebyshev filter, etc.

The proposed normalizations may include, without limitation, normalization with respect to the admission conditions, population-based normalization, and [-1,+1] normalization.

The proposed selection criteria may include, without limitation: domain-specific selection, physiologically-based selection, and clinically relevant selection.

The proposed integration techniques may include, without limitation: median, principal component analysis, independent component analysis, k-means clustering, fuzzy c-means clustering, artificial neural networks, multilayer perceptron, radial basis function network, support vector machines, decision trees, random forests, generalized linear models, and partial least square regression.

As such, the following system and method are capable of tracking the conditions of a patient, providing an individualized monitoring tool which makes use of population information, and can be tuned according to specific clinical needs.

The following also provides the capability to estimate, from the composite variability, the likelihood of developing a specific clinical outcome (i.e. a clinical likelihood). Such a clinical likelihood can be estimated by joining the information of the composite variability with the information from clinical factors (e.g. body temperature, Body Mass Index, etc…), and can be used to describe the conditions of the patient at the moment of

investigation. The technique used to estimate the clinical likelihood may be, for example, logistic regression.

The clinical outcomes that are addressed by the composite measures that may be generated according to the principles below include, without limitation: extubation outcome from mechanically assisted breathing, spontaneous breath trial outcome, development and degree of sepsis, the onset of septic shock, development and degree of multiple organ dysfunction syndrome, and the degree of physical exercise. As such, the following provides different types of composite measures of variability; those measures are tunable to the specific clinical needs of a patient, and can estimate the likelihood of developing specific clinical outcomes. Moreover, the composite variability and the clinical likelihood represent two modules, which can be provided alone or together with others inside a monitoring system.

Accordingly, the following principles enable: the combining of single organ and multiple organ variability measures into one composite measure; the tracking of a composite variability measure over time; the combining of variability measures into one composite measure reflecting the overall degree of change in the physiological conditions; the combining of variability measures into one composite measure correlating with specific clinical outcomes; the combining of variability measures into one composite measure sensitive to the physiological and pathological stimuli causing variability; the combining of composite measures of variability with clinical information to produce likelihood of development of specific clinical outcomes; the tuning of composite measure of variability to detect specific clinical outcomes, through the selection of a variety of normalization, selection and integration procedures applied to variability measures; the tuning of

composite measures of variability to investigate domains of variability, through the use of a variety of normalization, selection and integration procedures applied to variability measures; and the tuning of composite measures of variability to investigate the physiological phenomena underlying variability, through the use of a variety of normalization, selection and integration procedures applied to variability measures.

Turning now to the figures, Figure 5 illustrates an exemplary system 8 for obtaining multiple measures of variability (also variability measure) 14, and generating a composite measure of variability (also composite variability) 20 therefrom. In the example embodiment shown in Figure 5, several variability analysis components 12 are shown, each at a corresponding monitoring site 11. Each variability analysis component 12 comprises software, hardware, or a combination thereof for monitoring variability of physiological parameters associated with one or more organ systems and computing or otherwise generating variability measures 14. Figure 5 illustrates that the variability measures 14 may be generated and sent over a network 16, stored in a data store 18 for later use, and multiple variability measures may be obtained from the same variability analysis component 12 (see uppermost variability analysis component 12). It can be appreciated that the variability analysis component 12 may in general represent any server, computing system, clinical equipment, service, or other data component that is operable to execute computer executable instructions for generating the variability measures 14 using monitored physiological parameters. Further detail regarding how the variability measures 14 are computed and in which environments is provided in the next section.

The variability measures 14, whether sent over a network 16, stored in a data store 18 and transported using another medium, are provided in this example embodiment to a composite variability component 10. It can be appreciated that as shown in Figure 5, the composite variability component 10 may be coupled to, integrated with, or otherwise operable with a variability analysis component 12 and thus the separation between such components 10, 12 is shown for illustrative purposes only. The composite variability component 10 comprises software, hardware, or a combination thereof which is programmed or configured to utilize variability measures 14 to generate a composite variability 20. It can be appreciated that, similar to the variability analysis component 12, the composite variability component 10 may in general represent any server, computing system, clinical equipment, service, or other data component that is operable to execute computer executable instructions for generating a composite variability 20 using multiple variability measures 14.

As shown in Figure 5, the composite variability 20 may be stored in a data store 22, provided to a monitor 24 for display purposes, or provided to another system 28 over a network, to name a few.

The following section provides further detail regarding various examples of how the variability measures 14 may be obtained.

Obtaining Variability Measures

As discussed above, the composite measures of variability may be considered an extension of individual variability measures. As such, the generation of a composite measure of variability can be applied in any context in which variability measures, obtained from a variability analysis component 12, can be applied. For example,

composite variability measures can be generated using data obtained in real-time, older data, data obtained in an intensive care unit (ICU), data obtained using portable monitoring devices recording variability, etc. A composite measure of variability therefore is not dependent on any particular mechanism for obtaining the variability data, so long as a set of variability measures is available, a composite measure can be obtained, as explained in greater detail below. The following illustrates three exemplary monitoring sites 11 to demonstrate the various ways in which the variability measures 14 can be obtained in order to generate a composite variability measure 20. Further detail concerning an underlying software framework for obtaining and distributing variability data can be found in applicant's co-pending U.S. Patent Application No. 12/752,902, published under US 2010/0261977 to Seely, the entire contents of which are incorporated herein by reference.

An example of a hospital monitoring site 11a is shown in Figure 6. The elements shown in Figure 6 are meant to illustrate several possible components that may interact with one another at the hospital site 11a, however, any number (or all) of these elements can be used or not used in specific hospital sites 11a depending on the actual equipment and/or personnel present at the hospital site 11a and the needs of the patients 33 and personnel. In addition, the parameters being monitored (and the monitors themselves) may differ from network to network. As will be explained, at each monitoring site 11, including the hospital site 11a shown in Figure 6, is at least one variability analysis server 12' for using acquired data to conduct variability analyses over time and generate data files 30 that can be viewed at the site and provided to, for example, a central service (not shown). However, as shown, each variability analysis server 12' can interface with multiple

patients 33 and, as such, typically only one variability analysis server 12' is required at each monitoring site 11. The variability analysis server 12' gathers data acquired from one or more patients 33 through individual patient interfaces 34, computes the measures of variability (i.e. conducts variability analyses) for one or more patient parameters, and connects to, for example, a central server through the Internet, for facilitating the transfer and/or receipt of data files 30, threshold data 31 and update data 32. As shown, there can be different types of patients 33 such as those in the ICU or in a regular hospital ward.

The patient interfaces 34 monitor physiological parameters of the patient 33 using one or more sensors 35. The data or patient parameters can include any variable that can be accurately measured in real time or intermittently. The data may be obtained from a continuous waveform (at a certain frequency level, e.g. 100Hz for a CO2 capnograph or 500Hz for an EKG), or taken as absolute measurements at certain intervals, e.g. temperature measurements. The sensors 35 and patient interfaces 34 may include, for example, an electrocardiogram (ECG), a CO_2 capnograph, a temperature sensor, a proportional assist ventilator, an optoelectronic plethymography, a urometer, a pulmonary arterial catheter, an arterial line, an O_2 saturation device and others. To provide more meaning to the data acquired through the sensors 35, clinical events are associated with the data, through an act of recording time stamped events 36, which are typically entered by a heath care worker 37 in the hospital (bedside) environment. Clinical (time stamped) events can be physical activity, administration of medication, diagnoses, life support, washing, rolling over, blood aspiration etc. The clinical events are associated with a specific time, which is then also associated with the data that is acquired at the same specific time using the sensors 35. It will be appreciated that the clinical events can also

be recorded in an automated fashion, e.g. by utilizing algorithms which detect events electronically and process such events to designate them as clinical events or noise. In this example, the patient interface 34 is configured to gather the time stamped event data 36 concurrently with the sensor data 35, further detail being provided below. It may be noted that additional non-time-stamped information (e.g. demographics) can also be recorded for each patient.

As can be seen in Figure 6, the variability analysis server 12' not only connects to the patient interfaces 34 and the Internet, but also to several other components/entities within the hospital site 11a. For example, the server 12' can interface with a hospital monitoring system 39 such as a nurse's station, as well as a central monitoring and alert system 38. The central monitoring and alert system 38 is capable of monitoring the variability analyses performed by the variability analysis server 12' in order to detect critical or potentially critical situations evident from such variability analyses and provide an alert or alerts to a medical professional 42, who can also receive data directly from the variability analysis server 12'. The variability analysis server 12' can be embodied as a fixed unit or a moveable unit such as on a cart, in order to facilitate movement about the hospital site 11a to serve multiple patients 33 in multiple locations. Similarly, the variability analysis server 12' can be a proprietary apparatus or can be embodied as an add-on to existing beside or centralized equipment to minimize space.

The variability analysis server 12' can also interact with a bedside monitor 40, which may be made available to or otherwise represent a nurse or other personnel that monitors the patient 33 at the bedside. Similarly, the variability analysis server 12' can also interact with sensor displays 44, which are associated with other medical equipment such

as ECGs, blood pressure sensors, temperature sensors etc. As noted above, the variability analysis server 12' can be a separate, stand-alone unit but may also be integrated as a plug-in or additional module that in this case could be used or integrated with existing bedside monitoring equipment, displays and sensors. Figure 6 also shows other monitors 46 which can include any other monitoring system or equipment that either can provide useful medical data or patient data 48 or would benefit from the data acquired by the variability analysis server 12'. Patient data 48, e.g. provided by an electronic patient database (not shown) or manually entered can also interact with the variability analysis server 12'. As will be discussed below, the patient data 48 may be appended to, or included with the data files 30 to provide further context for the data contained therein. This enables patient specifics such as age, general health, sex etc. be linked to the acquired data to assist in organizing data into demographics. As also shown in Figure 6, the variability analysis server 12' can provide data or otherwise useful information for local scientists 50 that are interested in or involved in the implications and effects of variability. It will be appreciated that patient privacy and other concerns can be addressed as required, by adding data security or other de-identification measures.

Turning now to Figure 7, a clinic site 11b is shown. An example of a clinic site 11b is a bone marrow transplant clinic. Similar to the hospital site 11a discussed above, the clinic site 11b includes a variability analysis server 12', that obtains data from one or more patient interfaces 34, and connects to the Internet for facilitating data transfer (i.e. to send data files 30 and to receive threshold data 31 and update data 32). In the clinic site 11b, the patients 33 are referred to as outpatients as they are not admitted to a hospital. The sensors 35, clinical events recorded as time stamped events 36 and patient data 48 is

acquired and used in a manner similar to that discussed above and thus further details need not be reiterated. Similarly, the variability analysis server 12' can provide data and interact with medical professionals 42 at the clinic site 11b, as well as local scientists 50, if applicable. The clinic site 11b may include one or more variability analysis servers 12', and would typically include a monitoring centre 52 that monitors the analyses of the various outpatients 33 and provides alerts if necessary. The monitoring centre 52 enables the clinic's variability analysis server 12' to be monitored from a remote location and allows personnel to monitor several servers 12' if several are present in the clinic. In this way, a central monitoring centre 52 can be used to service several clinic sites 11b.

A mobile site 11c is shown in Figure 8. The mobile site 11c enables the capabilities of the variability analysis server 12' to be used outside of the hospital and clinical environments and, as such, in this embodiment, the mobile site 11c serves any "user" or "subject". For the sake of consistency, hereinafter the term "patient" will refer collectively to any user or subject. In this way, it may be appreciated that variability analyses can be performed on any user, including athletes, firefighters, police officers, or any other person that can benefit from monitoring variability of one or more physiological parameters. This can therefore extend to providing real-time monitoring in extreme environments such as during a fire, in a mine, during rescue missions etc. where variability can indicate a potentially critical situation. In all cases, variability can be monitored over time and analyzed on an individual basis for any patient 33 such that the resultant data is specific to that individual. Using the wider system 8 allows a central service to take advantage of the individual results for many patients 33 and ascertain further and more complete information. The mobile site 11c generally represents any site

that includes a variability analysis server 12', which connects to the system 8 and can communicate with one or more patients 33, whether they are patients in the traditional sense or another type of user.

In the example shown in Figure 8, the user 33 generally includes a mobile device 54 and has a number of sensors 35 that are in communication with a variability analysis server 12'. The mobile device 54 can also be used to provide inputs, e.g. for the time stamped event data 36, as well as to provide a display to the user 33 for entering parameters or to view display data 60 acquired by the sensors 35 and/or processed by the server 12'. The connections between the mobile device 54 and the server 12', as well as between the sensors 35 and patient interface 34 can be wired or wireless and the variability analysis server 12' can be a fixed unit at a base station or a portable unit such as on a cart at a monitoring centre. The mobile device 54 can be a personal digital assistant (PDA), mobile telephone, laptop computer, tablet computer, personal computer, or any other device that can provide an input device, a display and some form of connectivity for interacting with the variability analysis server 12', preferably in a completely mobile manner.

As noted above, each monitoring site 11 may include a variability analysis server 12'. Details of various embodiments of existing variability analysis apparatus and configurations can be found in U.S. Reissue Patent No. RE41,236 E to Seely, the entire contents of which are incorporated herein by reference.

Generating Composite Variability Measures

The origin and nature of the variability measures 14 will depend on the clinical application. The following describes how multiple variability measures 14 may be

integrated to generate a composite variability 20. A set of variability time series may be extracted through a windowed analysis of one or more physiological signals. For example, each 2.5 minutes, a 5 minutes window of those physiological signals is selected, and from it the variability measures 14 are computed.

The physiological signals can be extracted in real time from acquisition devices (e.g. electrocardiograph, Holter monitor) as well as taken from a database, and are submitted to the processing unit (e.g. variability analysis server 12') computing the variability algorithms. The algorithm for the composite variability 20 may then be implemented from these variability measures 14.

Figure 9 illustrates a block diagram of an example of a composite variability process 100, which includes various signal processing operations on the variability time series to obtain the composite variability 20. As seen in Figure 9, in this example embodiment, multiple variability measures 14 (e.g. variability measure 1 to variability measure N) are obtained and undergo a filtering stage 102. A normalization stage 104 is then applied to the filtered data, and a selection of a subset of variability parameters performed at stage 106. The outcome of the selection stage 106 then undergoes an integration stage 108 to arrive at the composite variability 20. Various examples of normalization processes are shown in Figure 9,which include, without limitation, admission condition 110, population based 112, [-1,+1], and custom normalization 116. Various examples of selection processes that may be applied in stage 106, are shown in Figure 9, which include, without limitation, a domain specific selection 118, a physiological-based selection 120, a clinically relevant selection 122, and a custom selection 124. Various examples of integration techniques are also shown in Figure 9, which include, without limitation, a

principal component analysis (PCA) technique 126, an independent component analysis (ICA) technique 128, a median technique 130, and a custom technique 132. Since PCA 126 and ICA 128 techniques create a number of time series equal to the number of time series provided to them, a component selection sub-stage 134 is also shown for enabling the selection of the component expressing the clinically relevant information. PCA 126 and ICA 128 are techniques allowing the representation of a dataset in another coordinate system. Therefore, the dimensionality of the problem remains unaltered (e.g. if PCA 126 is applied to 4 variability time series, 4 principal components are created). It can be appreciated that the different principal/independent components differ from each other, therefore the sub-stage 134 selection can be made according to for what the composite is being used. For example, in sepsis detection it has been found that the first principal component is the most correlated with the development of sepsis, therefore that principal component is selected. It may be noted that the first component is oriented in the direction of maximum variation, which provides a good index of change. It can therefore be seen that in addition to supporting various processes and techniques, the stages performed by the composite variability component 10 can support custom techniques and processes for adapting to different applications.

It can be appreciated that the filtering and selection stages 102, 106 may be optional in some applications. Although the normalization stage 104 is also optional, the multiple measures may not be comparable, thus diminishing the meaningfulness of the output.

An example set of graphs illustrating the processing of variability time series 140 to obtain a composite variability 20 is shown in Figure 10 through Figure 15. It can be appreciated that the composite variability processes described herein may include various

ones of the operations shown in Figure 10 to Figure 15, as well as other additional unspecified ones, since different applications may require different type of processing. In the example processing shown, Figure 10 illustrates multiple variability time series 140 (a, b, c, etc.). Time series (a) has been acquired using a standard deviation variability technique, time series (b) has been acquired using a Shannon Entropy technique, and time series (c) has been acquired using a Detrendend Flucutation Analysis Area Under the Curve (DFA AUC) technique. The variability time series 140 after undergoing the filtering stage 102 result in multiple corresponding filtered outputs 144 (a, b, c, etc.) as shown in Figure 11. The filtered outputs 144 are then subjected to the normalization stage 104 to obtain corresponding normalized outputs 148 (a, b, c, etc.) as shown in Figure 12. Figure 13 illustrates an integration 152 of the normalized outputs 148 thus providing a composite variability 20 in the graph shown in Figure 15.

Each of the stages shown in Figure 9 will now be described in greater detail.

In stage 102, each variability time series undergoes filtering through, e.g. a Savitzky-Golay first order filter to obtain the filtered outputs 144.

In the normalization stage 104, each filtered variability time series 144 undergoes a transformation that makes it comparable with the other filtered time series 144. Depending on the type of study, one or more of the following normalization techniques can be considered. In the following examples, consider a generic time series $x(n)$, where n relates to the n-th sample of the time series.

1. Normalization with respect to admission conditions 110: $x(n) = \frac{x(n)-\mu}{\sigma}$

In this technique, μ is the sample mean, and σ the sample standard deviation of the time series during a specific time after admission (e.g. from admission to 6 hours later). This normalization creates a relative measure, i.e. a measure which emphasizes the changes from a zero (in this case the admission conditions). Therefore, it measures the deterioration from a baseline of health. A possible use is the detection of sepsis after transplant. It may be noted that, it has been shown that patients after transplant start with healthy conditions (high variability) and deteriorate with time, causing a continuous drop in variability; this variability was shown to track the development of sepsis.

2. Population-based normalization 112

In this technique as shown in the graph 160 in Figure 14, for each variability measure, the population cut-offs are taken from the characterization of the variability in healthy 162, impaired 164, and critically ill patients 166. Then, as shown in Figure 16, the values are normalized in a scale from 0 to 10 using, for example, a sigmoid function. This normalization enables the direct association of variability to the degree of impairment, creating an absolute measure. Therefore, the emphasis is not on the changes from a starting point (zero) as in the first technique, but rather on the severity of illness. It may be noted that this normalization technique is possible because there exists a monotone relationship between certain measures of variability and the degree of impairment (three examples are provided in Figure 20).

3. Normalization between [-1,+1] 114, according to: $x(n) = 2 * \frac{x(n) - \max(x(n))}{\max(x(n)) - \min(x(n))} - 1$

When needed, this normalization technique can be used to provide signal processing techniques with a suitable range of values. Also, the [-1,+1] normalization can be used

for visualization purposes when there is the need to display more than one measure of variability. It may be necessary to bind the range of values to effectively display their dynamic into a monitor. The simultaneous representation of multiple measures could be useful for a qualitative assessment of the variability trends (each measure of variability is likely to behave in different ways over time).

When a wide array of variability measures is employed, it is possible to reduce their number through specific criteria applied during the feature selection stage 106. The criteria that may be used to select a subset of the variability parameters may include, without limitation:

1. Domain-specific selection 118: one or more of the available domains of variability are selected, neglecting the other domains. Possible domains are the scale-invariant, the frequency, the statistical, etc. This selection enables the investigation of the value of the individual domains of variability in clinical applications.

2. Physiologically-based selection 120: only the measures with a given physiological explanation are selected. An example is the selection of variability measures based on a frequential representation of the original physiological time series; those measures are indeed known to be related with brain activity. Therefore, this selection enables the association of physiological phenomena to clinical events.

3. Clinically relevant selection 122: only the measures that were found to be associated with a specific clinical outcome are selected. For instance, it has been found that certain variability measures continuously drops down from 36 hours before the development of septic shock. Therefore, the selection of those measures would lead to a composite variability 20 specific to the tracking of sepsis. Alternatively, a correlation-based

selection between each single measure and the expected trend in variability could be used. Only the measures with a correlation higher than a certain threshold would be selected.

It can therefore be appreciated that different types of normalization and feature selection approaches can be implemented, depending on the information of interest.

In order to incorporate multiple variability measures 14 into a composite variability 20, the integration stage 108 is performed. The integration stage 108 is used to fuse the variability measures 14 that survived the selection process 106 (if applied).

There are various possible techniques of integration, for example, the PCA technique 126, the ICA technique 128, and the median 130 of the variability. PCA 126 is a signal processing technique which captures the direction of maximum variance within a dataset, as well as the directions uncorrelated to the dataset. Therefore, PCA 126 allows the creation of a composite measure of variability 20 assessing the degree of change in variability. ICA 128 on the other hand can be used to find a set of independent signals generating the recorded variability. Therefore, ICA 128 is capable of identifying the underlying stimuli causing variability (e.g. circadian fluctuations, sympathetic/parasympathetic activity). The median 130, computed across the values of the measures of variability at a given time, gives a conservative measure of the entity of variation.

It may be noted that the type of integration used in stage 108, for example, PCA 126, ICA 128, or median 130, will depend on the intent of the application. Fo example, when the application intends on characterizing changes in variability that occurs as a result of a clear known time stamped intervention (e.g. SBT or sedation vacation) the ICA technique

128 is likely to be preferred, since the variability parameters that change at the time of the event can be selected, and a composite of those measures built. When there is ongoing monitoring and the application is attempting to detect something without an a priori known time point (e.g. an onset of infection), then the PCA technique 126 or the median technique 130 may be preferred. When one is trying to pick up a sensitive measure of a change in any variability measures 14, then PCA 126 is likely preferred. When one wants to detect a change in a large number of variability measures 14, then median 130 is likely preferred. Accordingly, the principles discussed herein provide a flexible, tunable, selective approach with various integration techniques for multiple clinical scenarios.

These integration methods, used in the present context, create a mathematical model of variability. It may be noted that the mathematical model can be based on the patient information, and therefore provides an individualized tool for the monitoring of the patient's condition. It can be appreciated that the mathematical model will vary based on the number of variability measures 14 used, the application, and the integration techniques used. For example, integration techniques such as PCA 126 and ICA 128 may apply a linear combination of the variability measures 14, i.e. multiplying each measure by a specific coefficient, and then summing the measures to obtain a unique time series.

From the integration stage 108, it is possible to create one or more composite measures of variability 20. For example, PCA 126 and ICA 128 technqiues create multiple principal and independent components. It is therefore possible to select only one of those components using the component selection sub-stage 134, and the selected components. Alternatively, it is possible to select one or more of those, and estimate the probability of

development of a specific clinical outcome (also referred to herein as clinical likelihood – discussed below).

An example procedure 170 to estimate the clinical likelihood 180 using a composite variability 20 is shown in Figure 15. The procedure 170 shown in Figure 15 can be applied to any clinical outcome detectable through variability analysis, and includes the following steps:

At step 174, the composite variability time series is/are normalized. At step 176, relevant features are extracted from the normalized time series (e.g. slope of the curves, sample mean, sample standard deviation, etc…). At step 178, a technique taken from the domain of Pattern Analysis is used to create the clinical likelihood 180, by using the composite variability, and joining to that information the one coming from several clinical factors 172 (e.g., body temperature, Body Mass Index, etc..). As illustrated in Figure 16, Logistic Regression is a Pattern Analysis technique suitable for this purpose. The Logistic Regression technique creates, through a linear model, a variable, which is z in Figure 16. The variable z is therefore transformed through a logistic function P(z), into a probability. For example, the following formula, also shown in Figure 16 may be used: $P(z) = \frac{1}{1+e^{-z}}$, where: $z = \alpha + \sum \beta_j X_j$. In Figure 16 it may be noted that $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are constants, while $\boldsymbol{X}$ is a vector made joining the composite variability features with the clinical factors. It may also be noted that Logistic regression is a deeply characterized technique widely used in clinical applications, providing statistical information on the estimated probability (e.g. confidence intervals, predictive intervals).

In the following, examples of applications of for a composite variability 20 are provided. The applications of the composite variability 20, and the clinical likelihood 180, include the scenarios discussed below, as well as various others. As such, the following examples are for illustrative purposes only. In each scenario presented below, it is assumed that the variability measures 14 are extracted from multiple physiological signals, from at least two organ systems.

A. Tracking of the development of sepsis

Consider a set of variability measures recorded from a patient that just underwent a bone marrow transplant. It has been shown, in the case of sepsis development, that the patient variability continuously drops down in the days following intervention. Using an admission condition normalization 110, without any particular selection, and integrating the information through a PCA technique 126, it may be possible to track the development of sepsis as shown in Figure 17. It may be noted that Figure 17 illustrates the tracking of sepsis centered at the time of sepsis detection (t=0).

It may be noted that the admission condition normalization 110 can enhance the change from a healthy baseline, and the PCA technique 126 can detect the direction of maximum variation of variability. This composite variability 20 would be specific for the detection of the occurrence of sepsis, promoting a faster and more effective care.

B. Estimation of the probability of extubation outcome

Consider a set of variability measures recorded from a patient that is under mechanical ventilation, and that will undergo a spontaneous breath trial (SBT) - a procedure used to

assess whether the patient can be successfully extubated). It has been found that there is a drop of variability during the SBT, and this quantity can be used to assess the probability of extubation outcome. Using the population normalization, selecting a set of measures known to be related with this clinical outcome (i.e. clinically relevant selection), and integrating the measures of variability with the median, a composite measure describing the degree of change during a SBT would be created as shown in Figure 18 and Figure 19 – graphs 194 and 196. It may be noted that Figure 18 and Figure 19 illustrate trends of the composite variability 20 for two patients monitored before and during an SBT trial. The graph 194 in Figure 18 corresponds to a failed extubation and the graph 196 in Figure 19 shows a passed extubation. At time 0, the SBT started. The composite variability 20 in this example was created using the population-based normalization 112, a clinically-relevant selection 122 based on previous research, and a median integration 130.

The combination of the composite with clinical factors (e.g. tidal volume, rapid shallow breathing index) in a logistic regression model would lead to a likelihood of successful extubation. This likelihood could be used by the physicians during the SBT to decide whether it is more valuable to keep the patient in the SBT phase, or to stop the SBT and either extubate or keep the patient intubated. Therefore, the patient's care would be improved and care costs would be diminished.

In Figure 20, examples of population trends of three measures of variability are shown. In Figure 20, a first graph 200 of a Shannon Entropy (shannEn) trend is shown; a second graph 202 of a Plotkin-Swamy energy operator (PSeo) is shown: a third graph 204 of a Fuzzy entropy (fuzEn) trend is shown. These measures were applied to eight

populations: subjects healthy and young (HY), healthy and old (HO), unhealthy and young (UY), unhealthy old (UO), patients that passed extubation before a spontaneous breath trial (pSBT), patients that failed extubation during a spontaneous breath trial (dSBT), patients six hours before the diagnosis of sepsis (6bS), patients six hours after the diagnosis of sepsis (6aS). As showed, there is a monotone relationship between the degree of impairment and the value of variability.

C. Sympathetic/Parasympathetic activity estimation

The sympathetic and parasympathetic activities of the brain regulate the heartbeat. Therefore, given the appropriate tools, it may be possible to indirectly measure these brain activities from heart rate variability (HRV). It has been found that through HRV it is possible to assess the degree of severity of Parkinson's disease. A composite measure of variability 20 useful for this purpose could make use of the [+1,-1] normalization 114 (needed to let the measures be comparable with each other), a selection of the measures of variability directly estimating the brain activity (i.e. physiologically-based selection 120) and the ICA integration technique 128, which would identify and select either the sympathetic or the parasympathetic component affecting variability. Some variability measures associated with parasympathetic activity are the power of an RR interval time series at high-frequency (0.15–0.40 Hz) and the SD2 index extracted from Poincaré plots; some associated to both sympathetic and parasympathetic activities are the low-frequency/high-frequency ratio, the SD1 index from the Poincaré plots and the Number of Variations from Symbolic dynamics.

This composite variability 20 could open the way to the creation of a set of clinically-oriented applications of variability for brain diseases, as well as represent a valuable

research tool to understand the physiology of the brain. Some studies using ICA for this purpose already exist, however they directly applied this method to RR interval time series, and not the variability measures extracted from it.

Industrial Applicability

The principles described herein provide a procedure to create a variety of clinically relevant time series, namely composite measures. Each composite measure describes in a continuous and individualized way the conditions of patient. Each composite measure is based on the continuous analysis of a variety of variability time series extracted from a variety of physiological signals. Each composite measure provides independent and clinically valuable information. Therefore, the methods described herein provide several measures that can be used alone or together, depending on the specific clinical outcome of interest, so as depending on the clinical needs of the patient.

Moreover, the principles discussed herein represent a further step along the path from a descriptive to a predictive medicine. The assessment at an early stage of the continuous deterioration of a patient's physiological condition would allow the prevention of further impairment. Better probabilistic prediction of the time at which a patient can be removed from e.g. a life support system can significantly decrease the costs of clinical care associated with certain critical conditions.

It will be appreciated that any module or component exemplified herein that executes instructions may include or otherwise have access to computer readable media such as storage media, computer storage media, or data storage devices (removable and/or non-removable) such as, for example, magnetic disks, optical disks, or tape. Computer storage media may include volatile and non-volatile, removable and non-removable

media implemented in any method or technology for storage of information, such as computer readable instructions, data structures, program modules, or other data. Examples of computer storage media include RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium which can be used to store the desired information and which can be accessed by an application, module, or both. Any such computer storage media may be part of the composite variability component 10, variability analysis component 12, etc., or accessible or connectable thereto. Any application or module herein described may be implemented using computer readable/executable instructions that may be stored or otherwise held by such computer readable media.

It will also be appreciated that the example embodiments and corresponding diagrams used herein are for illustrative purposes only. Different configurations and terminology can be used without departing from the principles expressed herein. For instance, components and modules can be added, deleted, modified, or arranged with differing connections without departing from these principles.

The steps or operations in the flow charts and diagrams described herein are just for example. There may be many variations to these steps or operations without departing from the spirit of the invention or inventions. For instance, the steps may be performed in a differing order, or steps may be added, deleted, or modified.

Although the invention has been described with reference to certain specific embodiments, various modifications thereof will be apparent to those skilled in the art

without departing from the spirit and scope of the invention as outlined in the claims appended hereto.

Claims:

1. A method comprising:

   obtaining a plurality of measures of variability; and combining the plurality of measures of variability to obtain a composite measure of variability.

2. The method of claim 1, wherein at least one of the plurality of measures of variability corresponds to a single organ.

3. The method of claim 1 or claim 2, wherein at least one of the plurality of measures of variability corresponds to multiple organs.

4. The method of any one of claims 1 to 3, further comprising tracking the composite measure of variability over time.

5. The method of any one of claims 1 to 4, wherein the composite measure of variability reflects an overall degree of change in at least one physiological condition.

6. The method of any one of claims 1 to 5, further comprising correlating the composite measure of variability with at least one clinical outcome.

7. The method of any one of claims 1 to 6, wherein the composite measure of variability is sensitive to physiological and pathological stimuli causing variability.

8. The method of any one of claims 1 to 7, further comprising obtaining a plurality composite measures of variability and combining the plurality composite measures of variability with clinical information to produce a likelihood of development of at least one clinical outcome.

9. The method of any one of claims 1 to 8, further comprising tuning the composite measure of variability to detect at least one clinical outcome by enabling selection of any one or more of normalization, selection, and integration procedures applied to the plurality of variability measures.

10. The method of any one of claims 1 to 8, further comprising tuning the composite measure of variability to investigate at least one physiological phenomenon underlying variability by enabling selection of any one or more of normalization,

selection, and integration procedures applied to the plurality of variability measures.

11. A computer readable medium comprising computer executable instructions for performing the method according to any one of claims 1 to 10.

12. A system comprising a processor coupled to a memory, the method comprising computer executable instructions for performing the method according to any one of claims 1 to 10.

Chapter 4

4 Physiology underlying variability

4.1 Commentary

The investigation of the nature of variability can occur from two main approaches: data-driven and model-based. Each one offers advantages and shortcomings. The former has a direct connection with real life, making it more appealing, yet prone to a variety of problems affecting real data, such as the presence of noise, artifacts, nonstationarities, and unknown confounding variables (i.e. those ones mainly accounting for inter-subject variability). The model-based approach does not have those problems, however being based on a set of assumptions, loses connection with the real world, thereby possibly undermining the applicability of the results in the real world.

The following manuscripts make use of both approaches to investigate the meaning of a set of measures of heart rate variability. In particular, using the model-based approach we tried to demonstrate that variability is able to track the changes in multiple physiological parameters, thereby supporting the view that the traditional association between variability and sympathovagal modulation is reductive and often inappropriate, as well as highlighting the need to perform multivariate analysis to understand the physiology underlying variability. On the other hand, we used the data-driven approach to show that the same set of measures behaves similarly during infection and exercise in two matched populations. Given that those two types of stressor have in common the effect on cardiac

regulation, this result highlights the limitations of univariate analysis in tracking phenomena outside of the pure cardiovascular physiology.

4.2 Manuscript 3: Independence of heart rate variability metrics and their joint ability to detect and track changes

To be submitted

My contribution: plan and execution of the analyses, writing of the manuscript, review and final approval.

Independence of heart rate variability metrics and their joint ability to detect and track changes

Andrea Bravi[1*], Christophe Herry[2], Andrew J.E. Seely[3], André Longtin[4]

[1]Department of Cellular and Molecular Medicine, University of Ottawa
[2]Ottawa Hospital Research Institute
[3]Ottawa Hospital Research Institute
[4]Department of Physics, University of Ottawa

[*]Corresponding author

***Abstract*— Heart rate fluctuations are characterized through multiple measures of variability. Although those measures can be classified into domains according to the type of information they extract, their physiological interpretation is challenging, especially because the degree independence among measures of variability remains unknown. In this context of limited understanding, the common approach is to select a small set of measures and evaluate them in specific clinical applications, leaving the potential value of simultaneous evaluation of multiple measures of variability to be determined. To investigate the value of multivariate variability**

analysis in silico, we used a physiologically-based mathematical model to synthetically generate R-R interval time series, and then study the behaviour of 102 measures of variability in the healthy case and in response to physiological intervention. The reported results characterize 1) the degree of independence between those measures (using a normalized distance based on mutual information), 2) their ability to sense a change in multiple physiological parameters (through statistical comparison of the median levels of variability), and 3) their ability to track changes to these physiological parameters over time (through correlation analysis). Our results demonstrate a high degree of independence among the measures, and that their ability to differentially detect and track changes in physiological parameters is enhanced when measures are combined. In conclusion, we showed that multivariate analysis of variability through a wide array of measures is not only recommended, but necessary if the final goal is to understand cardiovascular physiology, and translate that knowledge into clinical applications.

***Index Terms*— Composite measure of variability, heart rate variability, independent component analysis, multivariate time series analysis, mutual information**

Introduction

Heart rate variability (HRV) is the study of patterns of variation in an R-R interval time series, i.e. the series of time elapsed between two successive R-peaks in an electrocardiogram[22]. HRV is characterized by extracting multiple features called variability measures. Each measure represents a distinct means to reflect variation, and can be categorized in different domains (a domain defines the type of information the measure extracts – e.g. spectral, entropic, and scale-invariant[26]). In general, HRV has been documented to be sensitive to different stressors, either physiological or pathological[26], making them unspecific sensors (i.e. able to respond to heterogeneous

stimuli). However, because of the poor physiological understanding underlying variability[17,22], the choice of which measures to use for a given application is often arbitrary. Moreover, the interpretation of HRV becomes challenging due to the fact that measures coming from different domains of variability appear to be strongly correlated when using real patient data [1]. The quantification of the degree of interdependence between a set of variability measures could be addressed through many techniques, such as linear/nonlinear correlation, mutual information, independent component analysis[165], surrogate data hypothesis testing[166] and cluster analysis[167]. However, all these techniques make assumptions that can hardly be tested on real patient data because of multiple confounding factors, such as difference in demographics, clinical conditions and non-stationarity. A way to avoid the need to account for those confounding factors would be to have a synthetic dataset, where the differences between measures are exclusively linked to the independent information they bring. Therefore, this dataset would allow the identification of systematic correlation between different measures of variability.

To create this synthetic dataset, the present study takes advantage of a physiologically-based model by Kotani et al. [168] (2005). Kotani's model is a variation of Seidel and Herzel's model[2] presented in 1995, and further characterized years later [169–172]. As Kotani et al. highlighted, the main feature of his model is its ability to modulate the 1/f noise characterizing HRV, reproducing levels of scale-invariance qualitatively similar to the ones from real healthy and diseased patients.

Given the absence of confounding factors, Kotani's model represents a good tool to answer the following questions based on a comprehensive set (n=98) of measures of variability: 1) are those measures of variability independent, 2) are those measures sensitive to changes in multiple physiological parameters (i.e. not only sympathovagal modulation), 3) is it possible to track changes in a specific physiological parameter using variability analysis. Overall, these questions are meant to address the more pragmatic question: why use so many measures of HRV? The final goal is to understand what

independent pieces of physiological information are "captured" by each measure. This is important because clinicians are ultimately interested in physiological parameters, rather than abstract mathematical quantities. In the following, three different types of analysis are presented to investigate each of the three questions. Then, the results are shown and discussed, suggesting some answers on why using many measures of variability is important.

Methods

Model

Seidel and Herzel's model[2] is made of three compartments, namely the autonomous nervous system, the cardiovascular system and the respiratory system. More specifically, those compartments are described through a set of five differential equations (three of which have delays), respectively describing the dynamics of: 1) cardiac and 2) vascular norepinephrine, representative of the sympathetic nervous system's influence on the heart; 3) the phase of the heart, modelled as a pacemaker through an integrate-and-fire model; 4) blood pressure, affecting the activity of pressure sensors in blood vessels (baroreceptors); 5) the respiratory phase, affecting sympathetic and parasympathetic neuronal activity. The model is not stimulated by any source of noise. Because of the inability of that model to reproduce 1/f noise on HRV, Kotani et al. added two sources of noise: Brownian motion on a beat-to-beat basis on the baroreceptor activity, representative of mechanical noise from the vascular system, and 1/f noise on the parasympathetic activity. See the Appendix (Chapter 4.2.1) for a full list of the equations of the model.

Simulations and Variability

We reproduced the Kotani model, numerically integrating the delayed differential equations through Runge-Kutta method with a constant step size of 5ms and the use of ring buffers. Each time the simulation was run, the first 2 minutes were discarded to

exclude transients. The R-R interval time series generated by the model were then used to compute 98 measures of variability through a 5min windowed analysis (chosen based on the literature[22]) with no overlap, through the Continuous Individual Multiorgan Variability Monitoring (CIMVATM) software, developed with Matlab. The list and description of the measures is provided in an online manual[173]. Different window sizes (we tried 15 and 60 minutes) did not alter the results.

Datasets

We used five datasets for this work. **SIM_Health**, made of 1000 simulations of 7 minutes, making use of the physiological parameters (PAR) specified in Table 10, representative of a healthy subject[168].

Because the first two minutes of simulations were discarded, each realization allowed the computation of a single value on a unique 5min window for the 98 HRV measures; the values from all the realizations were thereby used to create 98 distributions of 1000 samples. Other three datasets were built through the same procedure, and making use of the PAR of Table 10, with the following exceptions. **SIM_SymVag** used $k^S_{c_{c\mathrm{Ne}}} = 0.84$ and $k^S_{c_{v\mathrm{Ne}}} = 0.6$, increasing sympathetic modulation, and $k_p = 0.22$, decreasing vagal modulation; these values were used by Kotani et al. to represent sympathovagal modulation in a subject with congestive heart failure[168].

Table 10 - Physiological parameters (PAR) representative of healthy subject

k_1	0.02	$\theta_{c\mathrm{Ne}}$	1.65
k_2	1.25 x 10^-3	$\tau_{v\mathrm{Ne}}$	2
$p^{(0)}$	50	$k^s_{c_{v\mathrm{Ne}}}$	0.5
$v_s^{(0)}$	0.95	k_v	0.2
k_s^b	0.8	$\theta_{v\mathrm{Ne}}$	4.2
k_s^r	3 x 10^-4	τ_{sys}	0.125
$v_p^{(0)}$	0	$S^{(0)}$	-13.8
k_p	1.1	k_s^c	10
k_p^b	0.036	k_s^t	45
k_p^r	4.5 x 10^-3	k_s^v	20
$T_{(0)}$	0.6	$\hat{S}$	70
$k_\phi^{c\mathrm{Ne}}$	1.6	n_s	2.5
$\hat{c}_{c\mathrm{Ne}}$	2	$\tau_v^{(0)}$	2.8
$n_{c\mathrm{Ne}}$	2	$\bar{\tau}_v$	1.2
k_ϕ^p	5.8	$\hat{c}_{v\mathrm{Ne}}$	1
$\hat{v}_p$	2.5	n_{vNe}	1.5
n_p	2	T_{resp}	3.5
θ_p	0.5	G	0.2
$\tau_{c\mathrm{Ne}}$	2	v_{trig}	1.3
$k^s_{c_{c\mathrm{Ne}}}$	0.7		

SIM_Baro used $k_s^b = 1.6$, expressing a reduction in the effect of the baroreceptor activity on the sympathetic activity. **SIM_Cont** used $k_s^t = 22.5$, expressing a reduction of the effect of the heart rate on the cardiac contractility. The last dataset, **SIM_Cont24**, is a single simulation of 24 hours, using a time varying k_s^t:

$$k_s^t = 45 + 3\sin\left(\frac{2\pi t}{\lambda}\right) \quad (1)$$

where t is the simulation time, and $\lambda = 100$ minutes. Because of the 5min windowed analysis with no overlap, this dataset was therefore made of 98 variability time series of 288 samples. The datasets are summarized in Table 11.

Table 11 - List of the datasets

Dataset name	Physiological description	Type of data	Altered PAR from Table
SIM_Health	Healthy subject	98 distributions	-
SIM_ SymVag	↑ Sympathetic activity ↓ Parasympathetic activity	98 distributions	$k_p = 0.22$, $k^s_{c_{c\mathrm{Ne}}} = 0.84$ $k^s_{c_{v\mathrm{Ne}}} = 0.6$
SIM_Baro	↓ Baroreceptor activity effect on sympathetic activity	98 distributions	$k^b_s = 1.6$
SIM_Cont	↓ Cardiac contractility due to heart contraction	98 distributions	$k^t_s = 22.5$
SIM_Cont24	Fluctuations in cardiac contractility due to heart contraction	98 time series	$k^t_s = 45 + 3\sin(\frac{2\pi t}{\lambda})$

$\lambda = 100$ minutes

Independence between measures

To investigate the level of shared information between measures of variability we used the normalized information distance (NID). NID is defined as

$$\mathrm{NID} = 1 - \frac{I(U,V)}{\max\{H(U),H(V)\}} \quad (2)$$

where $I(U,V)$ is the mutual information between two random variables U and V, and $H(\cdot)$ is their Shannon entropy[174]. We computed the mutual information using the kernel density estimation method[175], and the Shannon entropy using Scott's rule for the bin width[176]. Briefly, NID is defined in the range [0,1], where 1 means statistical independence between the two random variables (in our case between two measures of variability), and 0 means total information overlap. Given the absence of confounding

factors, two measures of variability were expected to show NID = 1 only when sensitive to different kind of information.

We evaluated the NID between the distributions of all the possible pairs of measures of variability (n=5151) on datasets **SIM_Health**, **SIM_SymVag**, **SIM_Baro** and **SIM_Cont**.

Sensitivity to changes in physiological parameters

The analyses described above investigate the degree of information shared between two measures, without addressing how each measure behaves with respect to changes in PAR. This is of particular interest to infer the behavior of the measures of variability with real physiological data. This was done through two analyses: 1) sensitivity to sympathovagal modulation was assessed by comparing variability from **SIM_Health** with the variability of **SIM_SymVag**; 2) sensitivity to changes in mechanical properties of the cardiovascular system was done comparing the variability of **SIM_Health** with the variability of **SIM_Baro**, and **SIM_Cont**. In particular, we used the Wilcoxon rank-sum test to check whether there was a statistical difference between the medians of either **SIM_SymVag**, **SIM_Baro** or **SIM_Cont**, compared to the medians of **SIM_Health**. Bonferroni correction was used to account for multiple comparisons (α-value = $0.05/(98\times3) = 1.7\times10^{-4}$).

How to track changes in specific physiological parameters

Consider a system characterized by multiple observed variables $\{v_j\}$, each one affected by a set of mutually independent stimuli $\{s_i\}$. Independent component analysis (ICA) is able to provide an approximation of the stimuli from the observed variables, by simply combining them in a linear way, i.e.

$$s_i = \sum a_{ij}\, v_j \qquad (3)$$

The identification of the parameters a_{ij} is based on the maximization of the statistical independence between the reconstructed stimuli (e.g. minimization of mutual information, or maximization of non-Gaussianity) [165]. We used **SIM_Cont24** to evaluate the ability of tracking the change in k_s^t (i.e. the stimulus) by single measures of variability. In particular, this ability was quantified by computing the Pearson's correlation coefficient between the time series of an HRV measure and the stimulus. Then we compared this tracking ability for combinations of measures obtained through the FastICA[165] algorithm.

Results

The first analysis we performed was aimed at identifying the level of information shared between two different measures of variability. Figure 21 shows matrixes depicting the pairwise NIDs, as well as the distributions of NIDs for four datasets. Each row and

column of the NID matrixes represents one of the 102 HRV measures. The order of the measures is arbitrary, and has no particular meaning. The NID matrixes appear symmetric (as expectable, because the NID is a symmetric measure), and mainly red (i.e. NID close to 1), denoting high degree of independence between the measures.

The same result is shown through the distributions of the NIDs. From a visual inspection, both pairwise NIDs and NIDs' distributions appear similar across the four datasets, highlighting that the changes in the parameters of the model did not strongly alter the information shared between the measures. Overall, the figure revealed a great degree of independence between the measures of variability. In particular, at least 72% of the pairs of measures showed a NID in the range [0.9, 1], while a minimum of 1.4% had an NID below 0.5. Among the pairs of measures with an NID below 0.5 (low independence), we found: mean heart rate and activity; standard deviation and coefficient of variation; square root of the mean squared difference of successive R-R interval (RMSSD) and Poincaré SD1 (details on each measure are online[173]).

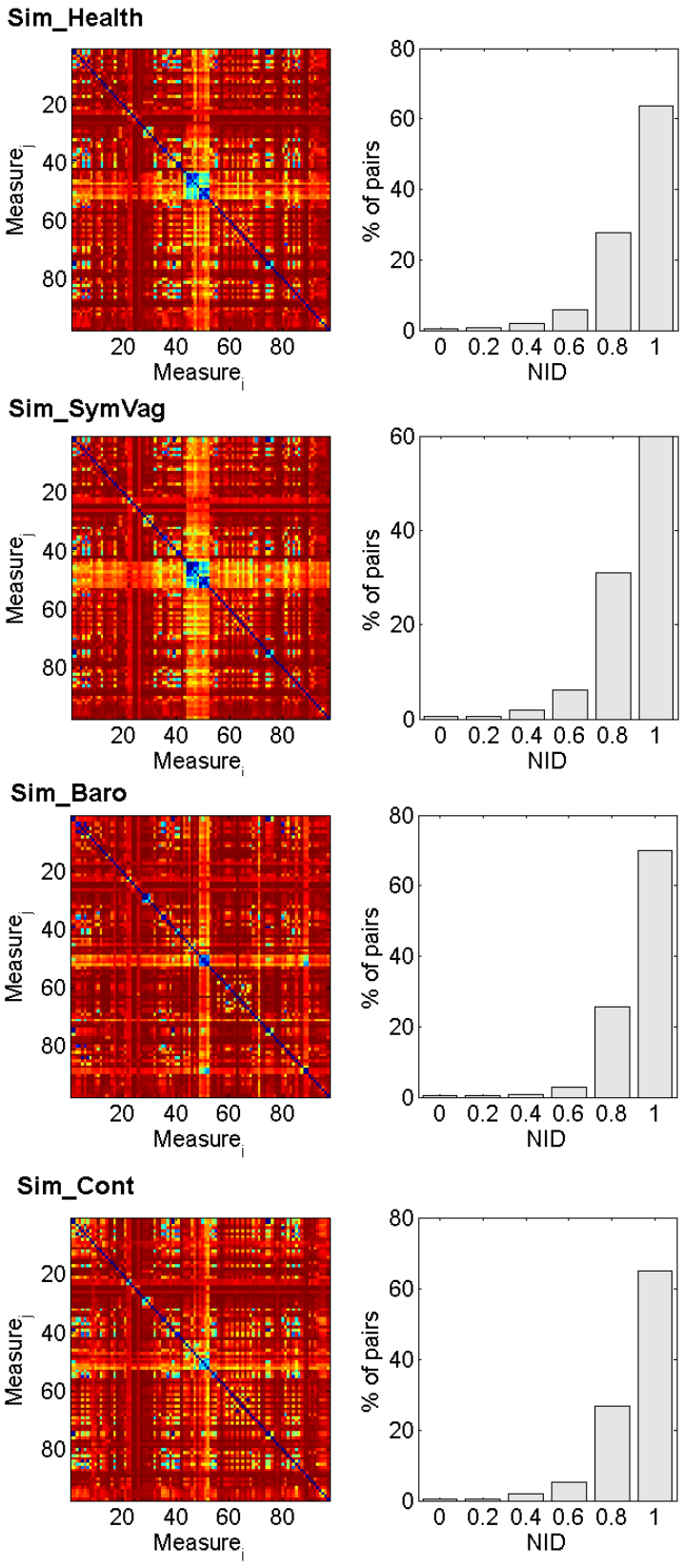


Figure 21- Normalized information distance (NID) analysis

The matrixes on the left depict the NID for all the possible pairs of 102 measures (please note that the order of the measures is random), for each of four datasets (Table 11). Notice that the diagonal with slope -1 represents the NID of a measure with itself, i.e. zero. The histograms on the right show the percentage of pairs within a given NID range. Each row of the matrix refers to a different dataset, as specified in the picture. In general, for each dataset at least 72% of the possible pairs were found to have a NID in the range [0.9, 1]. About 1.4% of the possible pairs had an NID below 0.5 for all four datasets.

The second analysis was aimed at investigating how the measures of variability behave when physiological parameters (i.e. PAR) change from a baseline of health. An example of response is provided in Figure 21, showing median and interquartile range of both mean heart rate and standard deviation of R-R interval time series computed with a set of PAR representative of an healthy subject (**SIM_Health**), and three alterations from that dataset – (see Table 11 for a summary).

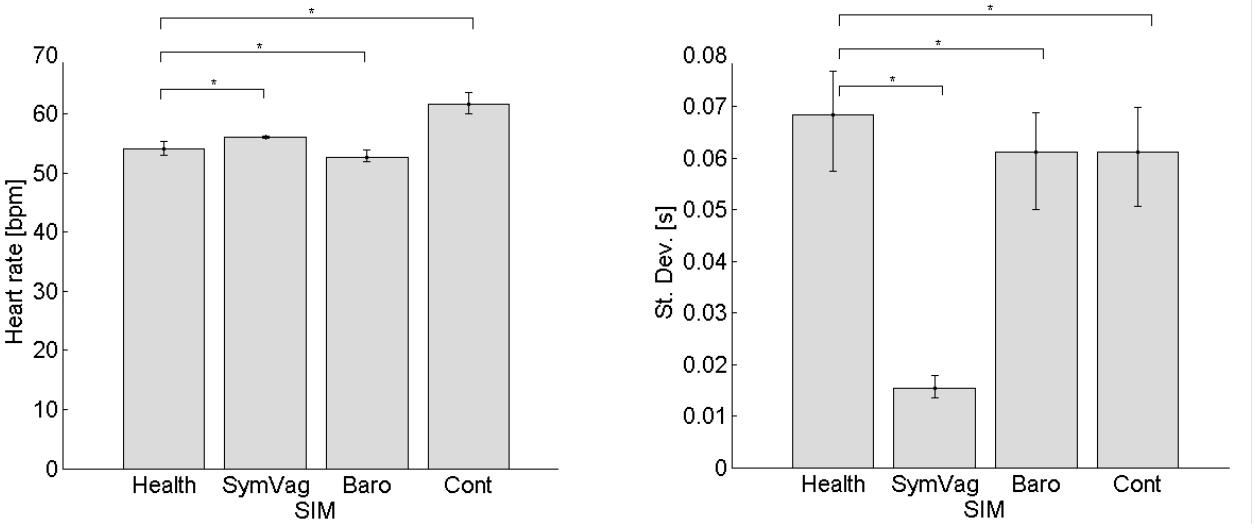


Figure 22 - Example of statistical comparison

The two bar graphs show the median and interquartile range of the mean heart rate (left) and the standard deviation (right) of R-R interval time series computed with a set of parameters representative of an healthy subject (SIM_Health), a dataset presenting sympathovagal alteration (SIM_SymVag), and two datasets showing alterations of the mechanical properties of the cardiovascular system (SIM_Baro and SIM_Cont) – (see Table 11 for a summary). The Wilcoxon rank-sum test was used to compare the difference in the medians between SIM_Health and the other datasets. The threshold for significance was 1.6×10^{-4} (Bonferroni corrected). All comparisons resulted statistically significant.

The summary of the behavior of the 102 measures of variability is reported in Table 12. A minority of the measures (n=8) did not detect any change from **SIM_Health**. Among

those that showed a change (n=98-8=90), only 14% (n=1+4+9=14) showed a statistically significant change in only one dataset. None of the domains of variability demonstrated a particular sensitivity to a specific dataset.

Table 12 – Response to changes in physiological parameters

Type of response			Number of measures showing the response (divided by domain)					Total
SymVag	**Baro**	**Cont**	Statistical	Geometric	Informational	Energetic	Invariant	
-	-	-	1	2	1		4	8
-	-	Δ	1					1
-	Δ	-	2	2				4
-	Δ	Δ	2	1			1	4
Δ	-	-	2	1		1	5	9
Δ	-	Δ	4	4	3	2	7	20
Δ	Δ	-	1	1	1			3
Δ	Δ	Δ	10	9	6	11	13	49
			23	20	11	14	30	98

This table shows how the different domains of variability behaved respect to (independent) changes in physiological parameters. In particular, we compared through Wilcoxon rank-sum test the median variability between one dataset created using parameters representative of an healthy subject (SIM_Health), and three dataset (SIM_SymVag, SIM_Baro and SIM_Cont) created by separately altering the parameters of Kotani's model. The description of these dataset is provided in Table 11. The symbol "Δ" was used to express a statistically significant rejection of the null hypothesis (difference in the median - i.e. detection of a change in the parameter), while the symbol "-" was used when no rejection occurred. A α-value of 1.7×10^{-4} was used to assess statistical significance. Based on its behaviour, each measure of variability was assigned to one of the $2^3 = 8$ possible types of response, each one corresponding to a row in the table.

The third analysis was aimed at tracking over time the alteration of a physiological parameter. The maximal Pearson's correlation between any of the 102 measure of variability and the stimulus was 0.6, and only 6% of the measures showed a correlation higher than 0.3. We evaluated all the independent components (IC) created by combining any two measures of variability, and selected the combination that correlated the most with the stimulus. Figure 23 shows the correlation between the stimulus, mean heart rate, standard deviation, and the independent component (IC) created by linearly combining

the latter two variables. IC presented a Pearson's correlation with the stimulus of 0.84, increasing the correlation between the mean heart rate and the stimulus by 40%.

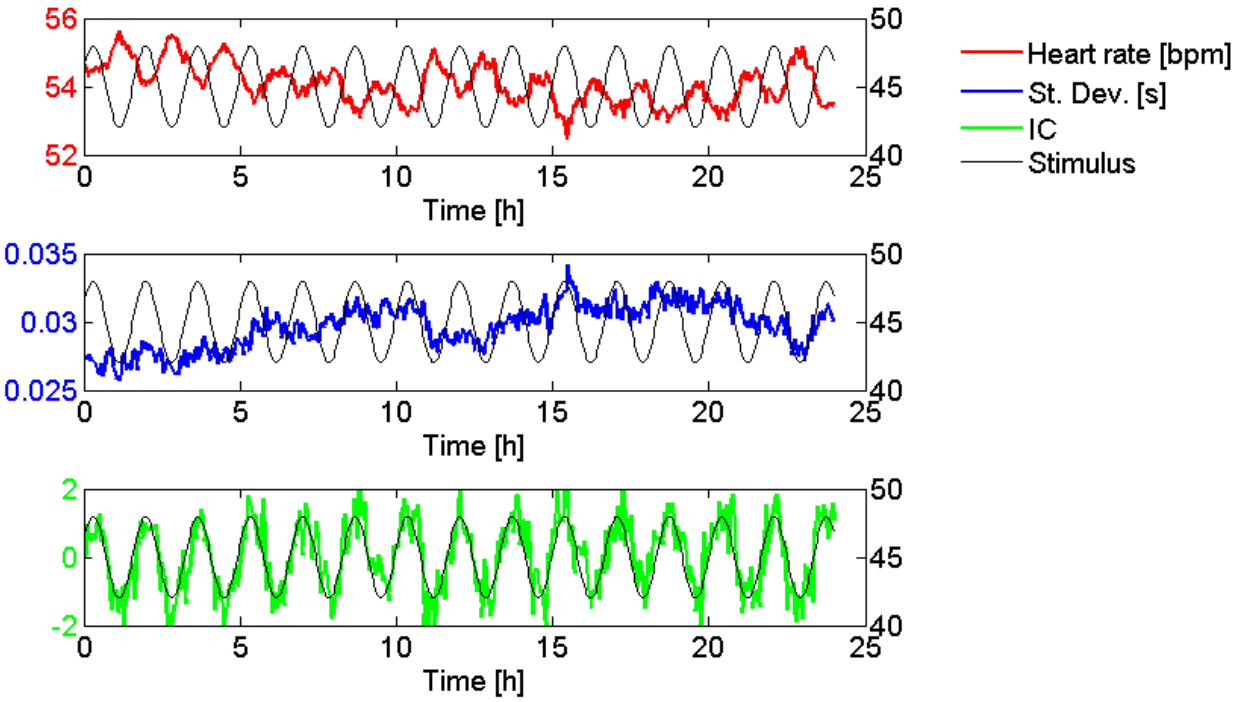


Figure 23 - Tracking of a physiological parameter
This figure shows how the mean heart rate, the standard deviation of the R-R interval time series, and a linear combination of the two made through independent component analysis (denoted with IC), track the fluctuations in $\boldsymbol{k_s^t}$ (namely, what we called a stimulus), over 24 hours of simulation (dataset **SIM_Cont24**). The Pearson's correlation between the stimulus and the other time series was respectively: 0.59 for the mean heart rate, 0.15 for the standard deviation, 0.84 for IC.

Discussion

Historically, HRV has been seen as a mean to measure sympatho-vagal modulation, with frequential measures such as LF and HF power as the primary means for its assessment. However, a recent opinion article highlighted how, over time, those measures gave inconsistent results in specific studies trying to highlight their physiological meaning [177,178]. The article highlighted how this observation was not limited to only frequential

measures, but also to scale invariant and entropic as well. Therefore, although the scientific community agrees on the clinical relevance of physiological variability [17,22,177], to date there is no clear understanding of what that means in physiological terms. A possible reason for these contradictions could be the fact that an HRV measure makes certain assumptions (e.g. stationarity, determinism, etc..) that are not respected by the R-R interval time series, therefore making the output of a measure unreliable. However, within studies (i.e. with a fixed protocol), there are several evidences of the fact that HRV measures show clinically relevant changes in a variety of applications [26]. Therefore this "noise" in the measure although present, may be not large enough to cover clinical differences. On the contrary, when the results of multiple studies are compared, inconsistencies appear. This seems to highlight that, on top of the inaccuracy of a measure, there may be a variety of confounding factors that affect its behavior and we do not take into account.

With the final aim of understanding which kind of physiological information a variability measure brings, the present article investigated 1) the degree of independence, 2) the ability to sense a change in multiple physiological parameters, 3) the ability to track over time changes in a specific physiological parameter, of a set of 102 measures of variability computed from synthetically generated R-R interval time series,.

With our first experiment we identified that the vast majority of the investigated measures under stationary conditions oscillate independently with each other (at least

98.6% of all the possible pairs had a NID higher than 0.5), meaning that they do actually capture different information. However, when a change in a physiological parameter occurs, most of the measures sense it and change accordingly, thereby producing correlation commonly observed with real patient data. The correlation that is found in a group of measures of time series is related to the complex nature of the cardiovascular system. Inherent to complex systems is their connectivity, and resulting ability to create behaviors that would not be achievable or explainable by separating the system into its multiple components. This connectivity makes a change in one physiological parameter affect the whole system, thereby affecting multiple measures of variability, albeit in different ways, as seen in our second experiment. For example, most of the measures of variability (n=9+20+3+49=81, see Table 12) detected a change in sympathovagal modulation. Those included the power at the low frequency (LF power), the power at the high frequency (HF power), and their ratio (LF/HF), namely measures known to be associated to sympathovagal modulation[22]. Furthermore, most of the measures (n=1+4+4+20+3+49=81) were able to detect changes in the mechanical properties of the cardiovascular system (i.e. show a change in either **SIM_Baro** or **SIM_Cont**). In most of the cases (n=4+20+3+49=75), the measures showed sensitivity to changes in at least two of the three datasets. We performed further analyses (data not shown), separately changing every parameter of Kotani's model, arbitrarily increasing and decreasing its value respect to the one representative of the healthy case (Table 10). Some of the trials led the system to unphysiologic behavior (e.g. constant concentration of norepinephrine), highlighting the need for a full systematic study of Kotani's model. Other showed trends

in the physiological variables qualitatively similar to the healthy case, yet confirming that no measure of variability was sensitive to exclusively one parameter.

These results highlight that **a measure of HRV can be seen as a nonspecific sensor of the cardiovascular system**, where "nonspecific" refers to the fact that it can sense multiple physiological mechanisms. This fact might provide an explanation about the inconsistent behavior that HRV measures show across studies, confirming the idea that there might be confounding factors we need to account for, to make sure different studies are comparable with each other.

Also, because in stationary conditions most HRV measures are independent with each other (even when computed within large windows of data), the number of physiological mechanisms each one is sensitive to, as well as the degree of sensitivity, are measure-specific. This means that **the physiology underlying variability could be better understood by trying to associate a physiological mechanism to a combination of measures of variability**. Indeed, by combining multiple measures that are sensitive to a given physiological mechanism, the "independent fluctuations" due to other mechanisms get "averaged out", while the common information gets strengthened. We saw this phenomenon with our third experiment, where it was possible to track changes in a physiological parameter by combining two measures into what we call a composite measure[179,180]. It is important to note that ICA is an unsupervised method, and therefore we did not force the independent components to behave like the sinusoidal stimulus in

any way. Also, in a previous study on real patient data, we created a composite measure to track sepsis development, showing that it was able to sense sepsis earlier and with more strength than the best performing single measure of variability[180]. While the idea of mixing multiple measures of variability to create predictive models is not new, to the best of our knowledge, this concept has never been applied to shed light on the physiological mechanisms behind variability, and provides a physiologically-sound explanation as of why a combination of measures is better than a single measure.

A major point of discussion is how much the *in silico* results can be extended to the *in vivo* world. Although Kotani et al. based their model on the physiologically-sound and well characterized model of Seidel and Herzel[169,171,172], their model still needs to be extensively characterized. Kotani qualitatively showed the ability of its model to reproduce the scale-invariant properties of the R-R interval time series, however no study has been made to quantitatively verify this result. Furthermore, other domains of variability were not investigated by Kotani, highlighting that particular care must be taken when analysing HRV measures not belonging to the energetic or invariant domain. Nevertheless, the fact that our results were confirmed in those two domains of variability, as well as all the others (Table 12), reassures us of their validity. Further efforts in the creation and validation of physiologically-sound models of the cardiovascular system will be useful to confirm our results and further investigate the relationship between physiological mechanisms and variability. In addition, the type of analyses we showed take advantage of the knowledge of physiological parameters that are not observable in the real world. Therefore, the identification of the correct way to combine measures of

variability into unique physiologically-based composite measures, is not straightforward. Also, although we recognize that multiple measures of variability are required because they bring independent information, their number needs to be kept to a manageable size for pragmatic reasons. In addition, it is certain that the true dimensionality of the HRV is likely to be considerably lower than the number of metrics to measure it. How to deal with these two problems is beyond the scope of this article, and is certainly of interest for future research. Nevertheless, the technical challenges associated with dealing with multiple measures of variability and real data do not undermine the principles highlighted in this article. Indeed, our results are in accordance with the literature, confirming the sensitivity of HRV to multiple physiological parameters (i.e. not only sympathovagal modulation), as expected by complex systems theory[17,22], and confirm what we observed when we tried to apply the concept of a composite measure in a pilot study with real patient data[180].

In conclusion, the key message of this article is that each measure of HRV is a nonspecific sensor of the cardiovascular system, bringing independent pieces of information about the state of the system. For this reason multivariate analysis of variability through a wide panel of measures is recommended if the final goal is to understand the physiology behind variability, and translate that knowledge into clinical applications. More specifically, because the measures of variability are sensitive to a wide array of physiological parameters, it is only by combining them into composite measures that further information can be exploited from the system.

4.2.1 Appendix

See Table 10 for the parameters' values.

Baroreceptor activity – v_b

$$v_b = k_1\left(p - p^{(0)}\right) + k_2\frac{dp}{dt} + \xi_1$$

Where ξ_1 is Brownian motion with standard deviation 0.16

Sympathetic activity – v_s

$$v_s^* = v_s^{(0)} - k_s^b\, v_b + k_s^r(1 - \mathrm{R})$$
$$v_s = v_s^*[\tanh\,(v_s^* \times 100) + 1\,]/2$$

Parasympathetic activity – v_p

$$v_p^* = k_p\left[v_p^{(0)} + k_p^b\, v_b + k_p^r(1 - \mathrm{R}) + \xi_2\right]$$
$$v_p = v_p^*\left[\tanh\left(v_p^* \times 100\right) + 1\,\right]/2$$

Where ξ_2 is 1/f noise with standard deviation 0.012.

Pacemaker phase – ϕ

$$\frac{d\phi}{dt} = \frac{f_s\, f_p}{T_{(0)}}$$

Sympathetic influence – f_s

$$f_s = 1 + k_\phi^{c\mathrm{Ne}}[c_{c\mathrm{Ne}} + (\hat{c}_{c\mathrm{Ne}} - c_{c\mathrm{Ne}}) \times \frac{(c_{c\mathrm{Ne}})^{n_{c\mathrm{Ne}}}}{(\hat{c}_{c\mathrm{Ne}})^{n_{c\mathrm{Ne}}} + (c_{c\mathrm{Ne}})^{n_{c\mathrm{Ne}}}}]$$

Parasympathetic influence – f_p

$$f_p = 1 - k_\phi^p\left[v_p(t-\theta_p) + \left(\hat{v}_p - v_p(t-\theta_p)\right) \times \frac{(v_p(t-\theta_p))^{n_p}}{\left(\hat{v}_p\right)^{n_p} + \left(v_p(t-\theta_p)\right)^{n_p}}\right]$$

Cardiac concentration of norepinephrine – $c_{c\mathrm{Ne}}$

$$\frac{dc_{c\mathrm{Ne}}}{dt} = -\frac{c_{c\mathrm{Ne}}}{\tau_{c\mathrm{Ne}}} + k_{c_{c\mathrm{Ne}}}^s\, v_s(t - \theta_{c\mathrm{Ne}})$$

Vascular concentration of norepinephrine – $c_{v\mathrm{Ne}}$

$$\frac{dc_{v\mathrm{Ne}}}{dt} = -\frac{c_{v\mathrm{Ne}}}{\tau_{v\mathrm{Ne}}} + k_{c_{v\mathrm{Ne}}}^s\,[v_s(t - \theta_{v\mathrm{Ne}}) + k_v]$$

Blood pressure – p

During diastole

$$\frac{dp}{dt} = -\frac{p}{\tau_v}$$

During systole

$$p = d_{i-1} + S_i\frac{t - t_{i-1}}{\tau_{sys}}\exp\left(1 - \frac{t - t_{i-1}}{\tau_{sys}}\right)$$

The transition from systole to diastole happens after τ_{sys} seconds from the time of contraction.

Cardiac contractility – S_i

$$S_i^* = S^{(0)} + k_s^c\, c_{c\mathrm{Ne}} + k_s^t\; T_{i-1} + k_s^v\, c_{v\mathrm{Ne}}$$
$$S_i = S_i^* + \left(\hat{S} - S_i^*\right)\frac{(S_i^*)^{n_s}}{\left(\hat{S}\right)^{n_p} + (S_i^*)^{n_s}}\Bigg]$$

Windkessel effect time constant – τ_v

$$\tau_v = \tau_v^{(0)} - \bar{\tau}_v[c_{v\mathrm{Ne}} + (\hat{c}_{v\mathrm{Ne}} - c_{v\mathrm{Ne}}) \times \frac{(c_{vNe})^{n_{vNe}}}{(\hat{c}_{vNe})^{n_{vNe}} + (c_{vNe})^{n_{vNe}}}]$$

Respiration – R

$$R = \cos(2\pi r)$$
$$\frac{dr^*}{dt} = \frac{1}{T_{resp}} - G \times (v_b - v_{trig})$$
$$\frac{dr}{dt} = \frac{dr^*}{dt}\left[\tanh\,\left(\frac{dr^*}{dt} \times 100\right) + 1\,\right]/2$$

4.3 Manuscript 4: Do physiological and pathological stresses produce different changes in heart rate variability?

Published in Frontiers in Physiology in date 30 July 2013.

My contribution: plan and execution of the analyses, writing of the manuscript, review and final approval.

Do physiological and pathological stresses produce different changes in heart rate variability?

Andrea Bravi[1,2*], Geoffrey Green[3], Christophe Herry[2], Heather E. Wright[4], André Longtin[5], Glen P. Kenny[4], Andrew J.E. Seely[1,2,3,6]

[1]Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario, Canada
[2]Ottawa Hospital Research Institute, Ottawa, Ontario, Canada
[3]Therapeutic Monitoring Systems Inc., Ottawa, Ontario, Canada
[4]Human and Environmental Physiology Research Unit, School of Human Kinetics, University of Ottawa, Ottawa, Ontario, Canada
[5]Department of Physics, University of Ottawa, Ottawa, Ontario, Canada
[6]Departments of Surgery and Critical Care Medicine, University of Ottawa, Ottawa, Ontario, Canada
[*]Corresponding author

Author contributions:

A.B., A.J.E.S. and A.L. conception and design of research; A.B. analyzed data; A.B., A.J.E.S, A.L., C.H., G.G., G.P.K. interpreted results; A.B., A.J.E.S, A.L., C.H., G.G., G.P.K., H.E.W. edited and revised manuscript; A.B., A.J.E.S, A.L.,C.C., G.G., G.P.K., H.E.W. approved final version of manuscript.

Abstract

Although physiological (e.g. exercise) and pathological (e.g. infection) stress affecting the cardiovascular system have both been documented to be associated with a reduction in overall heart rate variability (HRV), it remains unclear if loss of HRV is ubiquitously similar across different domains of variability analysis or if distinct patterns of altered HRV exist depending on the stressor. Using Continuous Individualized Multiorgan Variability Analysis (CIMVA^{TM}) software, heart rate (HR) and four selected measures of variability were measured over time (windowed analysis) from two datasets, a set (n=13) of patients who developed systemic infection (i.e. sepsis) after bone marrow transplant, and a matched set of healthy subjects undergoing physical exercise under controlled conditions. HR and the four HRV measures showed similar trends in both sepsis and exercise. The comparison through Wilcoxon sign-rank test of the levels of variability at baseline and during the stress (i.e. exercise or after days of sepsis development) showed similar changes, except for LF/HF, ratio of power at low and high frequencies (associated with sympathovagal modulation), which was affected by exercise but did not show any change during sepsis. Furthermore, HRV measures during sepsis showed a lower level of correlation with each other, as compared to HRV during exercise. In conclusion, this exploratory study highlights similar responses during both exercise and infection, with differences in terms of correlation and inter-subject fluctuations, whose physiologic significance merits further investigation.

Keywords: dimensions of variability, domains of variability, exercise, physical activity, disease, sepsis

Introduction

Several studies have shown that heart rate variability (HRV) can be used to characterize physiological (e.g. physical exercise, heat and cold stress), as well as pathological stress affecting the cardiovascular system. The utility of HRV in identifying illness states has been discussed in multiple reviews, and specific applications include the classification of heart failure, the prediction of mortality after myocardial infarction, and the estimation of autonomic modulation [26,15,154,22]. Similarly for physiological stress, measurement of HRV during physical exercise has been shown to estimate autonomic modulation, assess the level of fitness, characterize the beneficial effects of physical exercise, and many other applications [181–183].

One of the major findings extracted from this collection of results is that when the cardiovascular system is under stress, either physiological or pathological, there is a decrease in variability. When pathological stresses are considered, a widespread view is to interpret this loss as a decomplexification of the cardiac system due to illness [184,185]. In the domain of physiological exercise, however, a clear explanation for this phenomenon is lacking[181], nor is there a theory interpreting these two phenomena together. Developing a theoretical framework in this area of investigation might generate critical knowledge to understand the boundaries between health and disease, and allow knowledge translation from one domain to the other.

The first step to develop this framework consists of an extended comparison of how examples of pathological and physiological stresses affect HRV. With this in mind, the overall aim of this pilot investigation was to initiate, for the first time, a focused comparison of the patterns of change in HRV from a physiologic (i.e. exercise) and a pathologic (i.e. sepsis) stressor. The underlying hypothesis was that physiological and pathological stresses produce different effects on the body, and therefore their effects on HRV should differ. To test this hypothesis, we studied the changes in mean heart rate (HR), as well as four measures of variability belonging to different domains [26], and associated to distinct physiological dimensions [186]. Each domain of variability highlights the mathematical nature of a measure of variability, while the physiological dimensions are supposed to provide a physiological rationale to the changes in variability. Therefore, by investigating those four measures, we can shed light on the values of the two classification systems. In particular, we evaluated the trends (variation over time) of those measures from two datasets, one describing healthy subjects undergoing physical exercise under dry/warm conditions, the other ambulatory patients who developed sepsis after bone marrow transplant. Through analysis of the changes between baseline variability and variability during stress (i.e. either exercise or sepsis), as well as the analysis of the correlation between the measures, we highlighted similarities and differences between the two types of stressors.

Materials and Methods

Subjects

This study focuses on two datasets: 1) R-R interval (time between two successive R peaks during normal sinus rhythm) monitoring of 13 patients who developed systemic infection (i.e. sepsis) after bone marrow transplant (BMT), and 2) R-R interval monitoring of 13 matched (by gender, age, weight and height) healthy subjects undergoing controlled physical exercise under heat exposure, as described in more detail below. Table 13 summarizes the subject characteristics.

Table 13 – Patient Demographics

	Dataset		**p-value***
	EXERCISE (n=13)	**SEPSIS (n=13)**	
Gender • Male, n (%) • Female, n (%)	 9 (69%) 4 (31%)	 9 (69%) 4 (31%)	1
Age [years] – Median (95% CI)	50 (29 - 62)	49 (34 - 60)	0.83
Height [cm] – Median (95% CI)	169 (166 - 178)	173 (162 - 178)	0.97
Weight [kg] – Median (95% CI)	75 (62 - 90)	76 (63 - 105)	0.79
VO_2max [ml/kg/min] – Median (95% CI)	38.2 (31.5 - 43.2)	Not available	-

CI=Confidence Interval, * Wilcoxon rank-sum test was used to compare the medians, Chi-square test was used to compare the proportions

Sepsis

The first dataset (**SEPSIS**) consisted of patients undergoing BMT for hematological malignancy or other disorders. Inclusion criteria were treatment with myeloablative chemoradiotherapy followed by an allogeneic or autologous BMT. Exclusion criteria were pre-existing cardiopulmonary disease, taking beta-blockers or calcium-channel

blockers, pre-existing arrhythmia (e.g. atrial fibrillation, atrial bigeminy), and contraindication to electrocardiogram adhesives (e.g. allergy, severe psoriasis). Sepsis was defined as the systemic inflammatory response syndrome along with a clinically suspected infection requiring treatment. Over 50% of the patients were diagnosed with sepsis based on the presence of fever, defined a priori as one recording greater than 38.5 °C or two recordings greater than 38.0 °C within 12 h. The remaining diagnoses were based on clinical suspicion, bacteremia, productive cough and mucositis. For further details refer to [62,180]. The dataset consisted of 17 subjects: 3 who did not develop sepsis and showed an increase in HRV, 13 who developed sepsis and showed a reduction in HRV, and one insulin dependent diabetic subject who developed sepsis but showed an increase in HRV during the development of sepsis. To enable the comparison with the reduction in HRV during exercise, in this study we focused only on the 13 subjects who developed sepsis and showed a reduction in HRV. The average length of a recording was 12 days. A Zymed DigiTrak-Plus (Philips Healthcare, Canada) Holter system was used to record the R-R interval time series of each subject. Written informed consent was obtained from all participants, and the Ottawa Hospital Research Ethics Board authorized the study.

Exercise

The second dataset (**EXERCISE**) included healthy normally active (i.e. untrained but not sedentary) volunteers performing intermittent exercise in warm/dry conditions. We selected 13 subjects from a pool of 73, to match the septic patients. The matching was

done prior to any data analysis. The experimental protocol was approved by the University of Ottawa Health Sciences and Science Research Ethics Board. Prior to the experimental session, participants were asked to complete a Physical Activity Readiness Questionnaire (PAR-Q) to assess their eligibility to do physical activity. Maximal oxygen uptake (VO_{2max}) was subsequently measured during a progressive cycle ergometer protocol which consisted of a 2-min warm-up at 40 W followed by 20 W increments every minute until the participant could no longer maintain a pedaling cadence of at least 60 rpm.

The experiments were performed at the same time of day. Participants were asked to arrive at the laboratory after eating a small breakfast and to refrain from consuming alcohol and caffeine for 24 h prior to experimentation and to avoid major thermal stimuli on their way to the laboratory. Participants were also encouraged to arrive well hydrated as no fluid replacements were provided during the experiment. After 30 minutes of rest on a chair, participants performed four bouts of 15-min cycling (Corival recumbent cycle ergometer, Lode, Netherlands) at a constant rate of heat production (400 W) in an environmental chamber regulated at 35°C and 20% relative humidity. Each exercise bout was separated by 15-min of rest on the recumbent cycle, with a final 60 min recovery (not reported in the following analyses). Therefore, the length of each recording was of 150 minutes (60 of exercise, 90 of rest).

Electrocardiographic waveforms for each subject were recorded through a Holter DigiTrak XT (Philips Medical Systems, USA) physiological monitor. For simplicity we will refer to this dataset as **EXERCISE**.

Variability computation

For both datasets, only the beats characterized as normal sinus rhythm were included, while all premature beats were excluded. The classification was automatically performed by the Holter monitors and through delta detector on both datasets [187]). Using Continuous Individualized Multiorgan Variability Analysis (CIMVA^{TM}) software [173], HR and four measures of variability were extracted from the R-R interval time series of each subject through a windowed analysis, namely by using 5-minute windows for both **EXERCISE** and **SEPSIS**. This means that five values were computed using 5 minutes of data, and successive values of those measures were computed by repeatedly shifting the 5-minute window by 30 seconds for **EXERCISE** and 2.5 minutes for **SEPSIS**; these window steps were different because of the shorter length of the **EXERCISE** recordings compared to the **SEPSIS** recordings. This procedure created five variability time series per subject. Together with mean HR, the measures we investigated are standard deviation (statistical domain), the ratio between the power at low (LF) and high (HF) frequencies (LF/HF ratio - computed through the Lomb-Scargle periodogram, energetic domain), sample entropy (informational domain) and the Hurst exponent (computed through the Scaled windowed variance method, invariant domain). For details on the domains of variability, refer to Bravi et al. [26].

Statistical analysis

Before any analysis, each of the time series for the patients in the **SEPSIS** dataset were aligned with respect to the time of administration of antibiotics, and the time series for the subjects in the **EXERCISE** dataset were aligned with respect to the start of the exercise routine. Following the alignment, three analyses were performed. 1) Qualitative inspection of the population trends of HR and the four HRV measures (reported as median and 95% confidence intervals around the median). The population trends were computed by taking at each time instant the median value across the different subjects, as show in Figure 24.

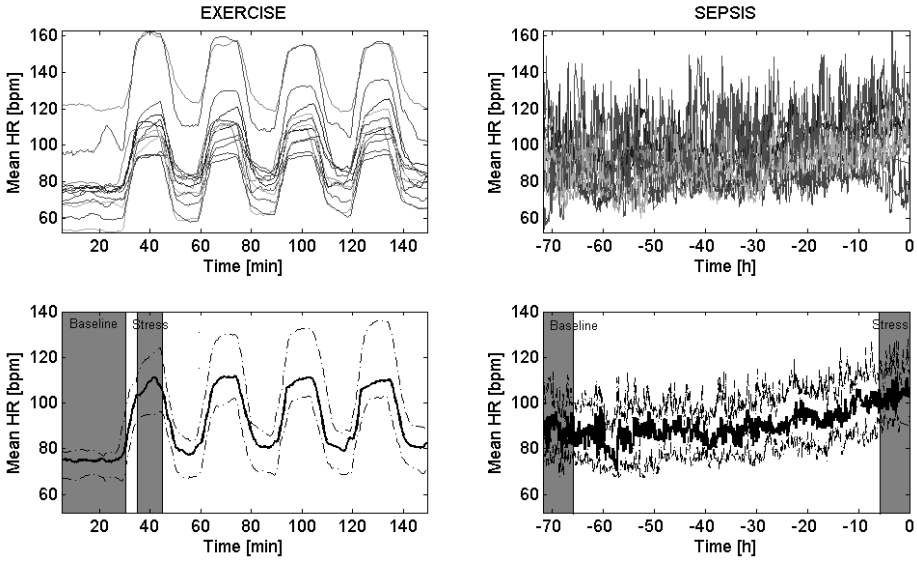


Figure 24 - Trends of mean heart rate in EXERCISE and SEPSIS
Top panels show how the mean heart rate (HR) changed over time for all the subjects (above), for either **EXERCISE** (left) or **SEPSIS** (right) datasets. Bottom panels report the median trends across each population – i.e. population trend - (bold solid line), together with the 95% confidence interval (dashed line). The **EXERCISE** time series start from 30 minutes prior the beginning of the first exercise bout, to the end of the last resting period. The **SEPSIS** time series start from 72 hours prior the administration of antibiotics, to the time of administration of antibiotics (t=0). In grey are highlighted the areas we referred to as "baseline" and "stress".

This led to five population trends for each dataset. 2) Statistical evaluation of the change between variability at baseline, and variability during stress. First, the variability time series were segmented, as summarized in Table 14, and then for each subject the median variability over time was computed, creating a distribution of values representative of either baseline or stress, for both **SEPSIS** and **EXERCISE**.

Table 14 - Segmentation of the time series

Dataset	Baseline			Stress		
	Start	**End**	**Length**	**Start**	**End**	**Length**
EXERCISE	30 minutes before the beginning of the first* exercise bout	At the beginning of the first* exercise bout	30 minutes	Ten minutes before the end of the first exercise bout	At the end of the first exercise bout	10 minutes
SEPSIS	72 hours prior the administration of antibiotics	66 hours prior the administration of antibiotics	6 hours	6 hours prior the administration of antibiotics	At the time of administration of antibiotics	6 hours

This table shows the time intervals that were used to compare the changes in HRV between baseline condition and stress. The variability time series were segmented as specified, and the median value in each time interval was computed. This procedure created two values (i.e. at baseline and during stress) for each subject and for each variability time series. Those intervals are reported in Figure 1 and 2 as grey areas.
* Different exercise bouts did not produce significant changes in the results

The Wilcoxon sign-rank test was used to evaluate the null hypothesis of zero change between baseline and stress. It is worth mentioning that what we defined as "variability at baseline" for the **SEPSIS** dataset is not really representative of baseline conditions, because likely some subjects were already developing sepsis 72 hours before the administration of antibiotics. However, not all the subjects had more than 72 hours of data prior the administration of antibiotics, therefore to include all of them we decided to use that time interval. 3) Spearman's nonlinear correlation analysis across the five population trends of **SEPSIS** and **EXERCISE** during the time intervals representative of "stress", as specified in Table 14.

Results

The alignment of the variability time series of each subject, as well as the creation of the population trends of the mean HR are reported in Figure 24. The figure shows that the mean HR considerably increased during exercise, and slightly increased during sepsis. Because of the low inter-subject variability, the 95% confidence intervals around the population trends are tight. Similarly, Figure 25 shows the population trends for standard deviation, LF/HF ratio, sample entropy and Hurst exponent. Standard deviation presented a clear population trend in both sepsis and exercise, both in terms of a decrease in time as well as a tightening of the confidence intervals.

The same phenomenon was not observed in either sample entropy nor Hurst exponent, even though they showed like standard deviation, a high sensitivity to nonstationarity (i.e. spikes in the transition between exercise and rest, and vice versa). Lastly, the LF/HF ratio showed a slight reduction in only the **EXERCISE** population trend, supported by a considerable reduction of the 95% confidence intervals around the population trends.

To evaluate the changes between baseline and stress in a quantitative way, the values of the measures (median and 95% confidence intervals), together with results of the Wilcoxon sign-rank test are reported in Table 15. Figure 26 provides a visual representation of the distributions of change from baseline to stress of Table 15.

Mean HR and standard deviation confirmed the changes observed from a visual inspection, rejecting the null hypothesis of no change in the transition from baseline to stress for both **EXERCISE** (p=0.0002) and **SEPSIS** (p=0.003). The LF/HF ratio rejected the null hypothesis only in **EXERCISE** (p=0.04). Hurst exponent and sample entropy were not able to reject the null hypothesis in neither **EXERCISE** nor **SEPSIS**.

Table 15 - Statistical results

	EXERCISE				**SEPSIS**			
Measure	**Baseline**	**Stress**	**Change**	**p-value**	**Baseline**	**Stress**	**Change**	**p-value**
Mean heart rate	74 (68, 78)	110 (95, 121)	36 (31 47)	**0.0002**	82 (79, 98)	99 (92, 114)	13 (-0.2 29)	**0.003**
Standard deviation	0.050 (0.026, 0.076)	0.014 (0.010, 0.021)	-0.031 (-0.059, -0.013)	**0.0002**	0.032 (0.023, 0.044)	0.016 (0.013, 0.028)	-0.015 (-0.035, -0.002)	**0.002**
LF/HF ratio	2.41 (0.99, 5.77)	1.09 (0.81, 2.36)	-0.63 (-2.31, 0.23)	**0.04**	2.94 (1.92, 6.18)	2.58 (1.50, 4.56)	-0.82 (-2.44, 1.61)	0.30
Sample entropy	1.67 (1.53, 1.94)	1.79 (1.17, 2.13)	0.08 (-0.21, 0.27)	0.73	1.57 (1.24, 1.71)	1.43 (1.25, 1.56)	-0.14 (-0.32, 0.30)	0.63
Hurst exponent	0.15 (0.08, 0.24)	0.13 (0.07, 0.24)	-0.01 (-0.09, 0.13)	0.79	0.21 (0.16, 0.25)	0.19 (0.14, 0.25)	-0.03 (-0.06, 0.07)	0.83

This table shows the values for each measure after the segmentation specified in Table 2, together with the Change (Stress minus Baseline, within each subject). Since baseline and stress were computed from repeated measurements on the same subjects, the p-values were computed through Wilcoxon sing-rank test. In bold are reported significant rejections of the null hypothesis of zero change (α-value=0.05)

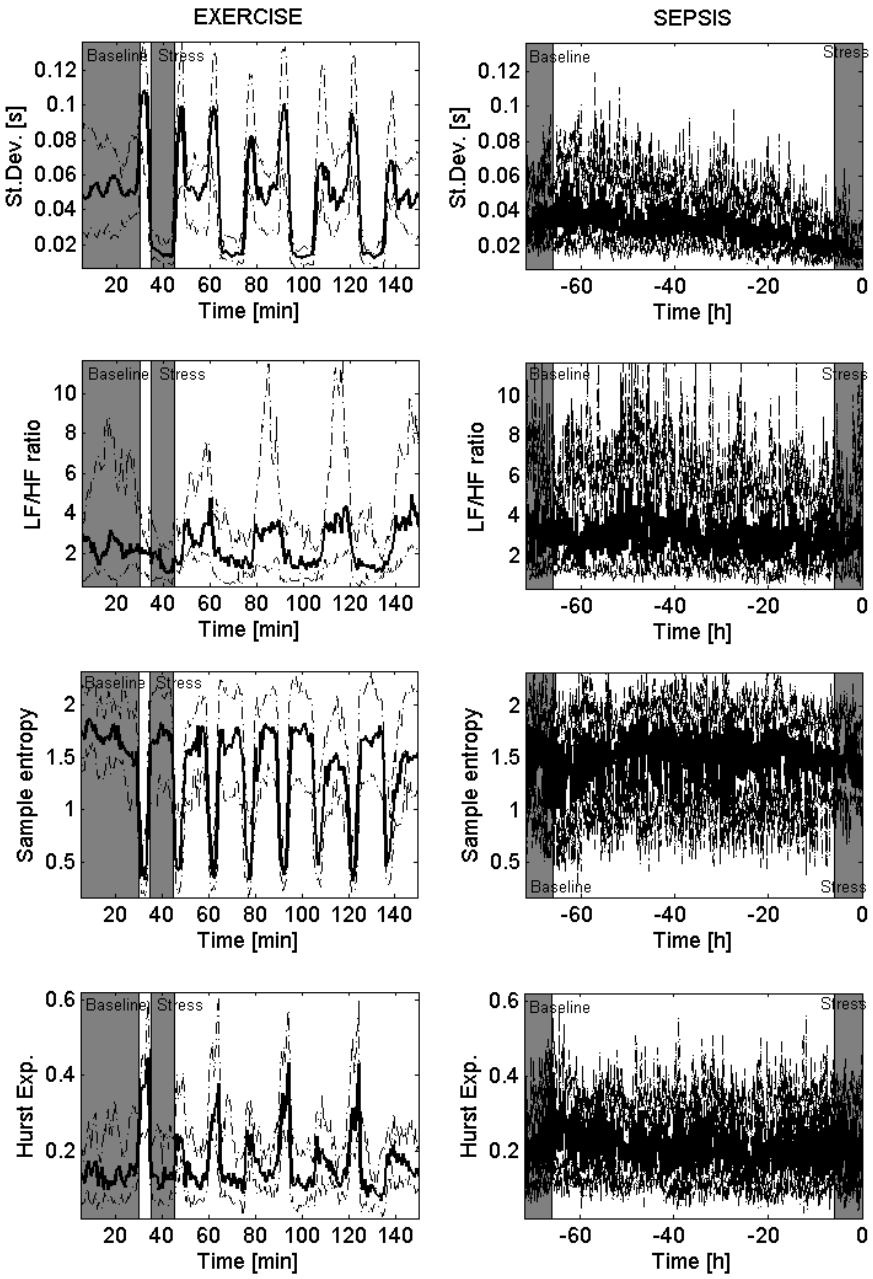


Figure 25 - Population trends in physiological dimensions of variability

These eight panels show the population trends of four measures of variability, hypothetically linked to separate physiological dimensions. From the top, each row presents: standard deviation, LF/HF ratio (computed through the Lomb-Scargle periodogram), sample entropy, and Hurst exponent (computed through Scaled windowed variance). The columns depict the median trend across all the subjects (bold solid line) and its 95% confidence intervals (dashed line) for the EXERCISE dataset (left) and the SEPSIS dataset (right). In grey are highlighted the areas we referred to as "baseline" and "stress".

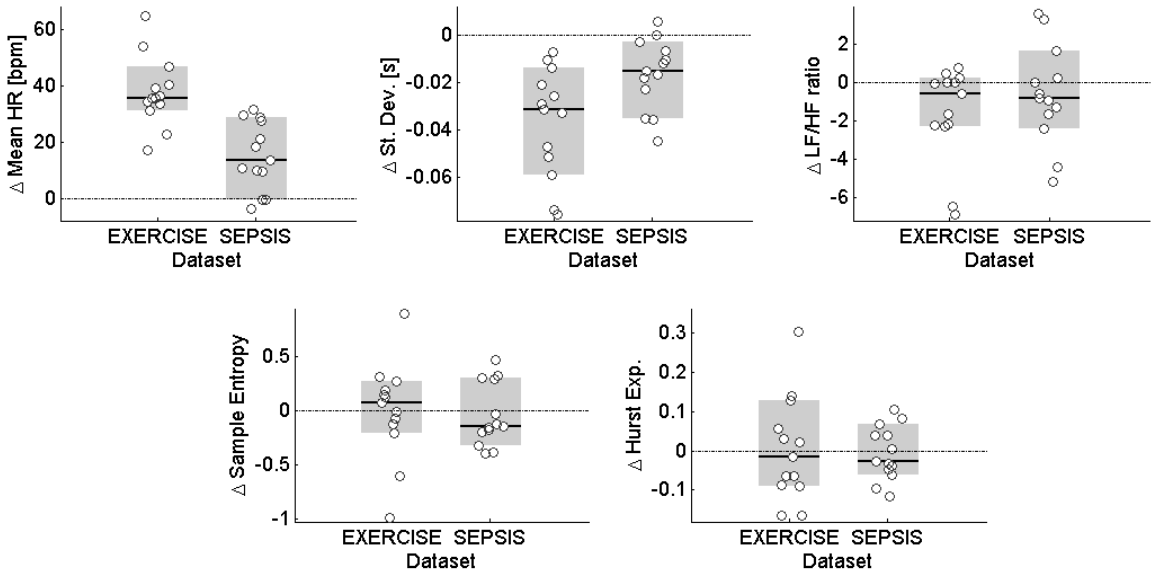


Figure 26 - Distributions of change from baseline to stress

The five panels show the distributions for mean HR and the four investigated measures of HRV. Each white circle represents the change in the median variability of a given subject from baseline to stress, as described in Table 15. The black solid line is the median of the dataset, and the grey area identifies the 95% confidence intervals for the median. The horizontal dash-dot lines highlight the lines of no change from baseline to stress.

The final analysis compared the degree of correlation between the population trends at the time of stress. The results are summarized in Table 16.

Table 16 - Spearman correlation matrix of the population trends in EXERCISE and SEPSIS, during stress

	Mean HR	Standard deviation	LF/HF ratio	Sample entropy	Hurst exponent
Mean HR	1.00, **1.00**				
Standard deviation	-0.68, **-0.40**	1.00, **1.00**			
LF/HF ratio	-0.76, **0.19**	0.71, **-0.17**	1.00, **1.00**		
Sample entropy	0.78, **-0.13**	-0.57, **0.05**	-0.58, **-0.14**	1.00, **1.00**	
Hurst exponent	-0.34, **-0.24**	0.22, **0.51**	-0.13, **-0.15**	-0.33, **-0.13**	1.00, **1.00**

This table shows the correlation values between the population time series in the two datasets, making use of only those samples representative of the stress phase (as specified in Table 14). The coefficients for **SEPSIS** are reported in bold, while the others are for **EXERCISE**.

Mean HR and standard deviation showed a good correlation (-0.68) during exercise, but lower correlation during sepsis (-0.40). Similarly, the levels of correlation between mean HR, standard deviation, LF/HF ratio and sample entropy resulted above |0.5| during exercise and below |0.5| during sepsis. The Hurst exponent showed low levels of correlation with all the measures in both sepsis and exercise, exception made for the correlation with standard deviation, which reached 0.51 during sepsis.

Discussion

The main objective of this study was to explore the differences between changes in variability due to exercise vs. sepsis; the underlying hypothesis being that the two types of stress produce differential impact on HRV measures.

Before interpreting the results, it is essential to highlight at a higher level what each measure of variability represents. It is well established that despite the differences from a mathematical point of view, many (but not all) measures of variability are highly correlated with each other [102], and are sensitive to different types of stressors affecting the cardiovascular system; this is compatible with our findings. Therefore, we believe that each measure of variability can be interpreted as a nonspecific sensor, which may be sensitive to multiple physiological mechanisms.

HRV demonstrated two main behaviors, possibly associated with two specific types of physiological phenomena: 1) sensitivity to exercise only (with “sensitivity” we intend the

ability to show a change between baseline and stress – as shown in Table 15), or 2) to both sepsis and exercise. Although a loss of complexity might have been expected with sepsis development as opposed to exercise, this was not observed. Furthermore, we repeated the same analyses described in this article with the 11 HRV measures found to be relevant in tracking sepsis development in a previous study making use of the same **SEPSIS** dataset [180], and we identified that all those measures changed significantly in both **SEPSIS** and **EXERCISE** (results not shown). This highlights that HRV responds similarly to both sepsis and exercise, although with differences, as exemplified by the particular behavior of the LF/HF ratio. This is similar to what was seen in a subsequent analysis on a set of other 17 measures of HRV, out of a pool of 96 measures (results not shown). Furthermore, when multiple measures of variability are compared through multivariate time series analysis (as done in Table 16 through correlation analysis), further differences between sepsis and exercise arise. This result supports the idea that the understanding of the physiology underlying variability may require the integration of information from multiple HRV measures.

The similarities between sepsis and exercise could be explained based on physiology. Exercise-induced immune responses and inflammatory-related immune responses share similar physiological mechanisms, such as the release of cytokines and mediators, albeit with a different intensity [188,189]. However, elevated levels of physical stress, such as those associated with performing exercise at very high intensity ($\geq$ 90% of VO_2peak), produce a much lower immune response compared to the septic one [188]. Because we saw major

changes during exercise rather than sepsis for both mean HR and standard deviation (Table 15), it is unlikely that immune response played a major role on the observed changes in HRV. Septic patients tend to show a normal or high cardiac output, due to an increased HR, a reduced afterload (due to the reduced vascular tone), and either increased (due to catecholamines) or reduced contractility (myocardial depression mediated by inflammatory cytokines) [190]. These similarities in the cardiovascular response could justify the similar behavior observed during sepsis and exercise, highlighting that major changes in HRV are likely driven by pure cardiovascular regulation.

On the other hand, the particular behavior of the LF/HF ratio is more difficult to explain, especially given the fact that this measure was documented to be sensitive to sepsis in multiple pilot studies, as specified in a recent review [191]. The key difference with the literature is that we did not compare septic patients with controls, but rather evaluated the change of LF/HF ratio during sepsis development. For instance, a study on 81 patients [192] showed that septic patients (n=21) who subsequently developed shock had lower LF/HF ratio respect to patients who did not develop sepsis (n=60). In our study the LF/HF ratio showed a reduction during sepsis development (Table 15) which was not statistically significant. However, the three patients which we excluded from the analysis because did not develop sepsis showed instead an average LF/HF ratio of 3.3 (CI: 3.0, 3.6) during stress, showing therefore a larger average LF/HF ratio (see Table 15), consistent with the literature. Although the lack of control of posture imperils our ability to further interpret the LF/HF ratio, it is of interest that this measure exhibited different behavior during

sepsis and exercise. For future reference, the Online Supplement 1 shows the details of LF and HF power (in normalized units) for both **EXERCISE** and **SEPSIS**, highlighting the major contribution of LF power in the LF/HF ratio population shown for **EXERCISE** (Figure 25).

An outcome of our analyses is that HRV measures taken from different domains of variability are likely to bring "unique" pieces of information, as shown by their low level of correlation (Table 16). The physiological interpretation of each measure in lights of the physiological dimensions of [186] is instead more challenging. Standard deviation decreased during both sepsis and exercise, possibly confirming the hypothesis that it represents the level of cardiopulmonary reserve of the body (i.e. the lowest the measure, the more the body is towards its maximal capacity). LF/HF ratio, measure of sympathovagal modulation, was reduced during exercise, rather than increasing as shown in [186]. Furthermore, despite the increase in HR, no significant change of LF/HF was detected during sepsis. Sample entropy, being a measure of complexity, was supposed to be primarily affected by a state of illness; despite the expectations, no change was observed during sepsis development. Lastly, the Hurst exponent has been linked to the capacity of the cardiorespiratory system to deliver oxygen and clear carbon dioxide. Given the increased demand of oxygen during exercise, we would have expected to see a decrease in the measure (H=0.5 indicates no fractality), which however did not appear.

Limitations

While a wider analysis using additional measures of variability is possible, we are cognizant of the small size of the patient populations, and wish not to over-analyze nor over-interpret the results. With a larger dataset in the future, it may be possible to better relate changes in HRV to the current theories about the nature of variability in physiological systems [32,186]. It is worth noticing also the need to characterize the sensitivity of each measure of variability to nonstationarity. As shown in Figure 25, standard deviation, sample entropy and Hurst exponent showed significant sensitivity to the transitions between rest and exercise (and vice versa). This was not a problem for the data analysis, given we knew the acquisition protocol; however, if we imagine the monitoring of ambulatory patients, the necessity of distinguishing nonstationarity from clinically relevant changes in HRV becomes clear.

An inherent limitation to comparing these two cohorts is the lack of standardization of the severity of the stressor, particularly true for the subjects developing sepsis. Indeed, most of the subjects in that group where diagnosed based on fever and clinical impression, possibly introducing a confounding factor in the analysis. Further work on the definition of what "stress" is and how to quantify it would enable a better comparison of different physiologic and pathologic forms of stress. On the same line, the choice of the specific type of physiologic and pathologic stress is a limiting factor of the analyses, and the investigation of alternative protocols or pathologies would be of interest to validate the results. Another limitation is that, the patients in the two populations were not matched

by fitness level (e.g. VO_2max), because that information was unavailable for the septic population. Given the similar changes seen in univariate HRV in both populations, it is likely that the fitness level did not produce a bias in the analysis. A similar argument could be applied to other systematic differences between the two datasets, e.g. different definition of baseline state or degree of intensity of the stressor. Also, our results are based on a set of arbitrary choices determined *a priori*, e.g. chosen window size, measures of variability, their computational parameters, time segments for the definition of baseline and stress (i.e. Table 14), and use of the median to create the population trend. A larger dataset would enable a sensitivity analysis respect to those parameters. Despite these limitations, this work represents an initial step towards the definition of the differences between physiological and pathological stressors, and the understanding of the physiological mechanisms underlying HRV.

Conclusions

Qualitative and quantitative comparisons of the changes in HRV during physiological and pathological stress were performed on two datasets, one from patients who developed sepsis after a bone marrow transplant (pathological stress), the other from a matched group of healthy subjects performing physical exercise (physiological stress). The comparison showed that HRV measures behave similarly during both exercise and sepsis development, although with subtle differences, whose explanation remains fertile ground for further investigation.

Conflict of interests

Andrew J.E. Seely is founder and Chief Science Officer of Therapeutic Monitoring Systems, Inc. (TMS), created to commercialize patented Continuous Individualized Multi-organ Variability Analysis (CIMVA) technology, with the objective of delivering variability-directed clinical decision support to improve quality and efficiency of care. Geoffrey C. Green is currently employed by TMS in the position of Product Manager.

Acknowledgments

We gratefully acknowledge the support of the Ottawa Hospital Research Institute (OHRI), the Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council (NSERC) and MITACS for this research program. The authors would like to acknowledge Joanie Larose for her role in the data collection of the **EXERCISE** dataset.

Chapter 5

5 Clinical applications of variability

5.1 Commentary

The creation of clinical decision support systems (CDSS) is a challenging task, requiring expertise from multiple fields. Indeed, the technical difficulties that were discussed in the introduction of this thesis represent only half of the problems associated with CDSS design. The aspect that has received comparably less attention and is of growing interest in the last years is the sociotechnical component of the system, i.e. the social context in which the technology will be used[193–195]. In particular, whether the context is an alarm, diagnostic or management system, for the success in clinical practice of a CDSS it is essential to take into account: 1) how the system positions itself in the existing clinical routine, 2) how the users interact with the system, 3) how environmental factors affect this interaction. A proper design of those three factors, united with the appropriate performance, leads to the construction of a relationship of trust between the user and the CDSS, thereby increasing its efficacy. When this relationship of trust is not built because of a wrong design of the CDSS, not only may the users start ignoring it, but the system may become source of distraction or irritation, thus actually producing an adverse effect on care.

The following two manuscripts present two CDSS entirely based on physiological variability. The first manuscript describes a system for the identification and monitoring

of sepsis development. The rationale behind the system is to provide clinicians with two pieces of information: 1) a continuous trend related to the degree of stress of the cardiovascular system, which the clinician can use as any other monitoring tool, and 2) an alarm system with high specificity and sensitivity, to be used to provide an alarm only when the estimated probability of developing sepsis is above 99%, thereby reducing false alarms. The second manuscript describes a system for extubation outcome prediction. This system takes data recorded during standard-of-care tests for extubation outcome prediction, called spontaneous breathing trials, and provides to the clinician the fold-increase in risk of requiring reintubation following extubation (i.e. failing extubation). The fold-increase in risk is a statistical quantity most clinicians are familiar with, and translates to "how many times, with respect to the known average risk of extubation failure, a given patient is more likely to fail". The manuscript will show the increased efficacy of this measure over clinical parameters commonly used by clinicians, as well as its complementarity with respect to clinical judgment.

5.2 Manuscript 5: Monitoring and identification of sepsis development through a composite measure of heart rate variability

Published in Plos One in date 19 September 2012.

My contribution: plan and execution of the analysis, writing of the manuscript, review and final approval.

Monitoring and Identification of Sepsis Development through a Composite Measure of Heart Rate Variability

Authors and Affiliations
Andrea Bravi[1*], Geoffrey Green[5], André Longtin[2], Andrew J.E. Seely[1,3,4,5]

[1]Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario, Canada
[2]Department of Physics, University of Ottawa, Ottawa, Ontario, Canada
[3]Division of Thoracic Surgery, University of Ottawa, Ottawa, Ontario, Canada
[4]Department of Critical Care Medicine, University of Ottawa, Ottawa, Ontario, Canada
[5]Therapeutic Monitoring Systems Inc., Ottawa, Ontario, Canada

[*]Corresponding author

Abstract

Tracking the physiological conditions of a patient developing infection is of utmost importance to provide optimal care at an early stage. This work presents a procedure to integrate multiple measures of heart rate variability into a unique measure for the tracking of sepsis development. An early warning system is used to illustrate its potential clinical value. The study involved 17 adults (age median 51 (interquartile range 46-62)) who experienced a period of neutropenia following chemoradiotherapy and bone marrow transplant; 14 developed sepsis, and 3 did not. A comprehensive panel (N=92) of variability measures was calculated for 5 min-windows throughout the period of monitoring (12±4 days). Variability measures underwent filtering and two steps of data reduction with the objective of enhancing the information related to the greatest degree of change. The proposed composite measure was capable of tracking the development of sepsis in 12 out of 14 patients. Simulating a real-time monitoring setting, the sum of the energy over the very low frequency range of the composite measure was used to classify the probability of developing sepsis. The composite revealed information about the onset of sepsis about 60 hours (median value) before of sepsis diagnosis. In a real monitoring setting this quicker detection time would be associated to increased efficacy in the treatment of sepsis, therefore highlighting the potential clinical utility of a composite measure of variability.

Introduction

Tracking the physiological conditions of a patient is of utmost importance in a clinical setting. Treatments provided at an early stage of development of disease are indeed more likely to be effective, and the effectiveness is related to higher chances of survival and lower healthcare costs. Considering the development of severe sepsis, a retrospective study on 2,731 subjects showed that each hour of delay in the initiation of effective antimicrobial therapy is associated with a mean decrease in survival of 7.6% [196]. Furthermore, severe sepsis and septic shock are the most common causes of mortality in critically ill patients, with a mortality of approximately 50% and an average annual cost of $16.7 billion in the USA [197].

Despite the intensive research for the management of severe sepsis and septic shock, there is the lack of a tool capable to continuously monitor its development. In the domain of neonatal sepsis identification, Moorman et al. proposed the Heart Rate Characteristic [198,199], a logistic regression model combining variability measures applied to R-R interval time. Their approach demonstrated a remarkable reduction in all cause infant mortality in a 3000 patient randomized controlled trial [58], and is leading to commercial applications. However, stepping from neonatal to adult monitoring the scenario changes. Infection remains a clinical diagnosis confirmed in a delayed and insensitive matter by blood cultures, still the gold standard [200,201]. We first reported on continuous heart rate variability (HRV) measurements during the onset and resolution of clinically diagnosed infection in immuno-compromised ambulatory patients, finding altered HRV in different

HRV metrics occurring ~ 24-40 hours in advance [62]. In this study we re-analyze the data using a novel method to integrate altered variability.

The present article proposes a novel method, a composite measure of variability, to be used for the identification and tracking of sepsis development. The composite measure was created by applying a sequence of signal processing steps designed to enhance the change from a baseline of health, and integrate the clinically relevant information collected from 92 measures of variability. This broad number of measures was used to maximize the probability to detect clinically useful information from the R-R interval time series. These steps produced a unique, composite measure, which has potential to enrich clinical monitoring.

Materials and Methods

Dataset

The study included 21 ambulatory outpatients (age median 51 (interquartile range 46-62)) who underwent bone marrow transplant (BMT) for hematological malignancy or other disorders. Sepsis was defined as systemic inflammatory response syndrome along with clinically suspected infection requiring treatment. Over 50% of the patients had sepsis diagnosed based on the presence of fever, defined a priori as one recording greater than 38.5 degrees centigrade or two recordings greater than 38.0 degrees centigrade within 12 hours. Inclusion criteria were treatment with myeloablative chemoradiotherapy followed by an allogeneic or autologous BMT, and informed consent. Exclusion criteria were pre-existing cardiopulmonary disease, taking beta-blockers or calcium-channel blockers, pre-

existing arrhythmia (e.g. atrial fibrillation, atrial bigeminy), contraindication to electrocardiogram adhesives (e.g. allergy, severe psoriasis). Continuous Holter ECG data was collected (average 12 [SD 4] days of monitoring) for all patients in the study, starting approximately 24 h before their BMT and continuing through neutropenia until its resolution or until withdrawal from the study. The used Holter system, a Zymed DigiTrak-Plus (Philips Healthcare, Markham, Ontario, Canada), sampled the ECG at 175Hz with 10-bit amplitude resolution, and annotated all normal QRS peaks and arrhythmias, including premature atrial and ventricular beats. Only the beats that characterized normal sinus rhythm (NSR) were included, while all premature beats were excluded. RR intervals were derived from R wave annotations. Among the 21 patients, four patients dropped out within 24 h of initiation of monitoring due to discomfort or other reasons, leaving 17 subjects for analysis, 14 of which developed sepsis. Sepsis was defined as systemic inflammatory response syndrome along with clinically suspected infection requiring treatment. Written informed consent was obtained from all participants, and the Ottawa Hospital Research Ethics Board authorized the study. For further details refer to [62].

Signal Processing

Through a windowed analysis (5 minutes window size, 2.5 minutes overlap) of the RR interval time series, 92 variability time series were extracted for each subject. From now on we will refer to the variability time series with the word “measures”, for simplicity. All the subjects who developed sepsis, developed it after 6 days after admission (median value). Therefore, to reduce the fluctuations with time scales shorter than the time scale

of sepsis development, the measures were filtered through a Savitzky-Golay zeroth-order filter (length of 577 samples, i.e. ~ 24 hour). Given their considerable number, two data reduction steps were applied: one with the aim of selecting only the relevant information for this specific application, and the other with the aim of reducing the redundant information.

In the first step, the Spearman correlation coefficient (SCC) between the measures and a prototype function representing the expected trend during the development of sepsis was computed. The type of prototype function was arbitrarily chosen as a straight line going from +1 at admission time to -1 to the time of administration of antibiotics. The set of values [-1.+1] is arbitrary, because the SCC is a nonlinear operator which compares only the order between the values of the line and one measure of variability. Those values were specified only to highlight that the correlation between this line and a variability measure is positive only when there is a monotone negative relationship between the two (i.e. the measure is decreasing over time). The values of correlation were then bootstrapped 1000 times to get an estimate of the average population correlation for each measure. Then, 11 measures with the highest average correlation were selected (the number of measures that were selected is arbitrary, and due to keep only ~10% of the available measures).

In the second step of data reduction, the 11 survived measures (per patient) were processed through Principal Component Analysis (PCA) after admission condition

normalization. This normalization transforms the value of the measures into a percentage of change with respect to the first 24 hours after admission, according to the formula ΔHRV= [current – baseline]/range, where "baseline" is the mean variability for the first 24 hours after admission, and "range" is the maximum variability less the minimum variability within the same time frame. PCA is separately applied to each patient, taking the set of 11 measures, and computing the loading coefficients of the first principal component, which is the component oriented in the direction of maximum variance of the dataset [167]. Recalling that the loading coefficients represent the components of a vector, it is possible to create a unique population reference coordinate system by taking the median for each set of loading coefficients across the population. Given the population reference system, the projection of all the measures for each subject is computed, obtaining one time series per subject, which we call a composite measure of variability, or "composite". To make the model more robust, the population medians of the loading coefficients were computed by bootstrapping 1000 times their distributions.

To motivate the choice of creating a composite measure, for each subject we compared the changes in the composite composite variability with the relative changes (i.e. after admission condition normalization) of the single measures composing it.

To create an alert system identifying when a patient is developing sepsis, the information obtained from the composite measure time series was further reduced by extracting the sum of the energy of the series at very low frequencies, which represent those frequencies with the time scale of sepsis development (i.e. days). This summed energy, which we call

E_s for simplicity, was computed in the interval (0,6.3] µHz. This interval was selected to include the peak frequency characterizing each composite measure (see Figure 27).

The energy was extracted by computing the Discrete Fourier Transform of the composite and summing the absolute value of the transform in the specified frequential interval.

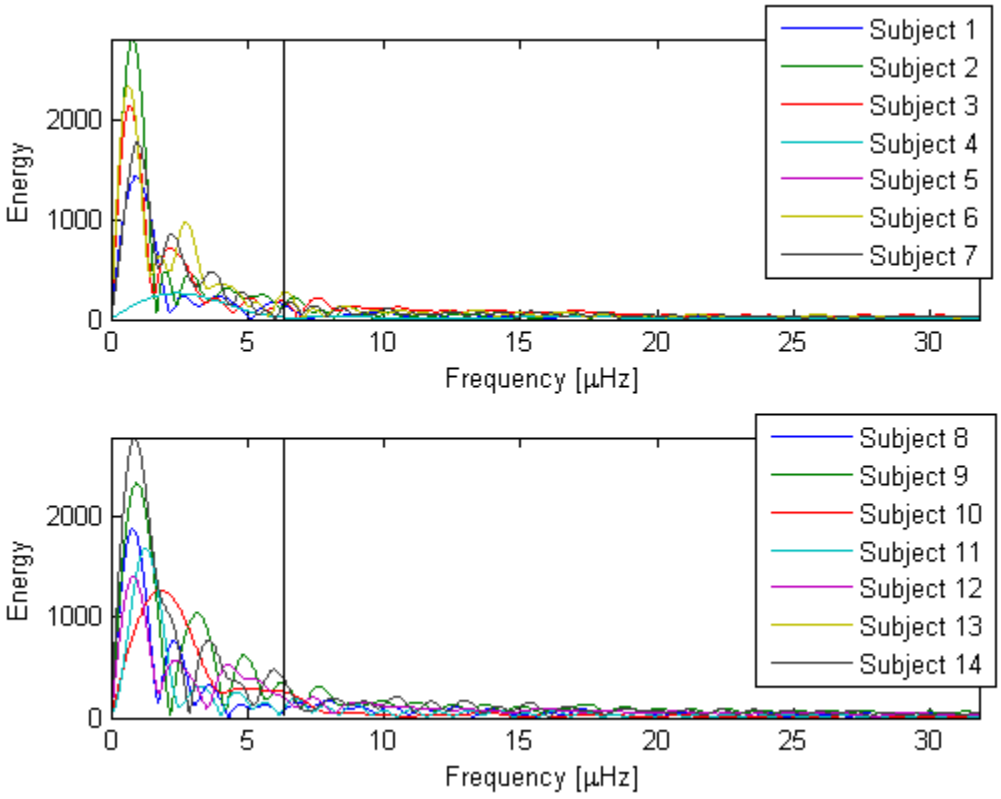


Figure 27 - Fourier transform of the composite measure
The figure shows the Fourier transform of the composite measure of variability for each subject. The black vertical line represents the threshold of 6.3 µHz, which was arbitrarly chosen to include the majority of the energy of the signals (i.e. the peaks in the transform). The energy at the very low frequency E_s was defined as the sum of the energy in the frequency interval (0, 6.3] µHz. A different selection of the threshold did not produce any change in the results, as long as those peaks were included.

Approaching the alert system creation from a classification perspective, two classes were introduced; the first one is the E_s during the second day after admission (we remind that the first day was used to pursue the admission condition normalization), namely E_{s1}, the second one is the E_s from the second day from admission to the moment of administration of antibiotics, namely E_{s2}. Then, in the training phase a logistic regression was used to create the decision boundary between the two classes, providing the probability of being developing sepsis; only the subjects who developed sepsis were considered in this phase.

Using a leave-one-out cross-validation, we tested the decision boundary for both subjects who developed and did not develop sepsis. To reproduce a standard monitoring situation, the probability of developing sepsis was assessed continuously over time by computing E_s over increasingly long intervals. This means that the same RR interval time series was analyzed taking multiple incremental windows (e.g. the probability is computed by taking the first 30 minutes of data, then the first 60 minutes, then 90, etc…) and reproducing the processing steps described above. To better simulate a monitoring situation, a new value of the probability was computed every 2.5 minutes (i.e. the window step of the windowed analysis). A diagram of the entire processing is reported in Figure 28.

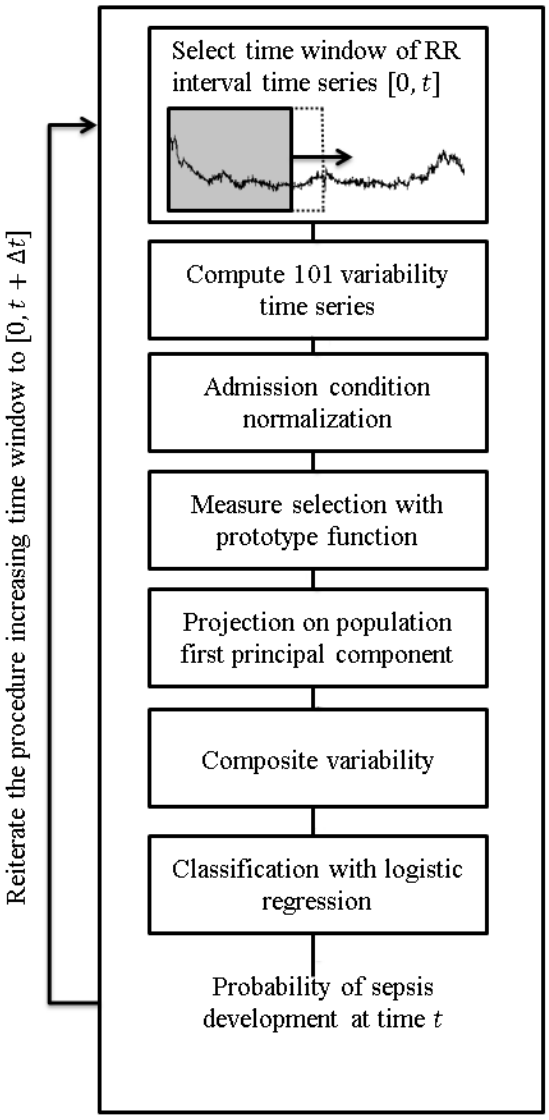


Figure 28 - Signal processing diagram
Block diagram showing how to create the composite measure of variability and the likelihood of developing sepsis. The time window [**0**,***t***] is increased at every iteration of 2.5 minutes. This allows to reproduce a monitoring situation where new R-R intervals are continuously analyzed. Having the variability up to at a certain time ***t***, we can compute the composite, and from the composite the probability of developing sepsis.

Because the E_s is a feature that increases either when variability is decreasing or increasing with frequencies lower than 6.3 μHz, we imposed the probability to be zero when the tested composite measure from which E_s is extracted was increasing or its last value in time was positive (the composite was considered increasing if the SCC with a line going from -1 to +1 was higher than zero).

This approach, based on the idea that only a decrease in variability corresponds to physiological impairment, is justified by the fact that the composite measures were found positive and increasing for the subjects who did not develop sepsis (Figure 29), and negative and decreasing when sepsis was developed (Figure 30) (see also Discussion).

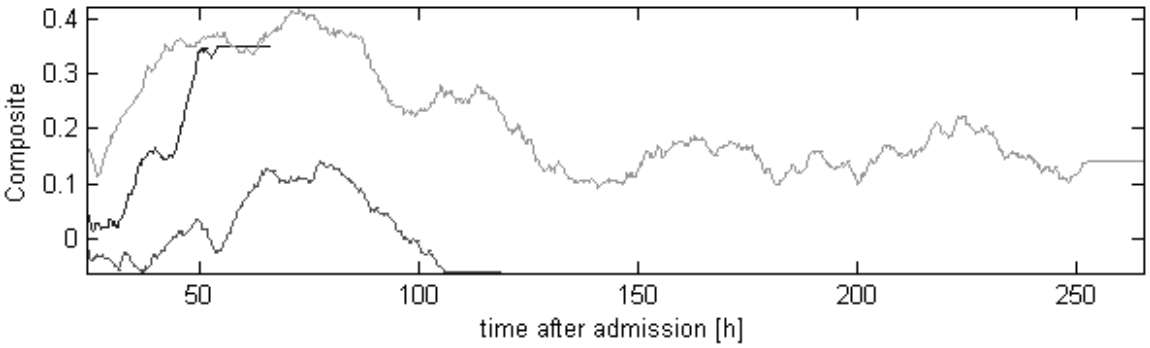


Figure 29 - Composite of subjects who did not develop sepsis

These time series represent the composite measures computed for the three subjects who did not develop sepsis. Because the admission condition normalization takes out the first day of recording, the composite starts from 24 hours after admission. The constant value at the end of the composite is due to the transient of the Savitzky-Golay filter.

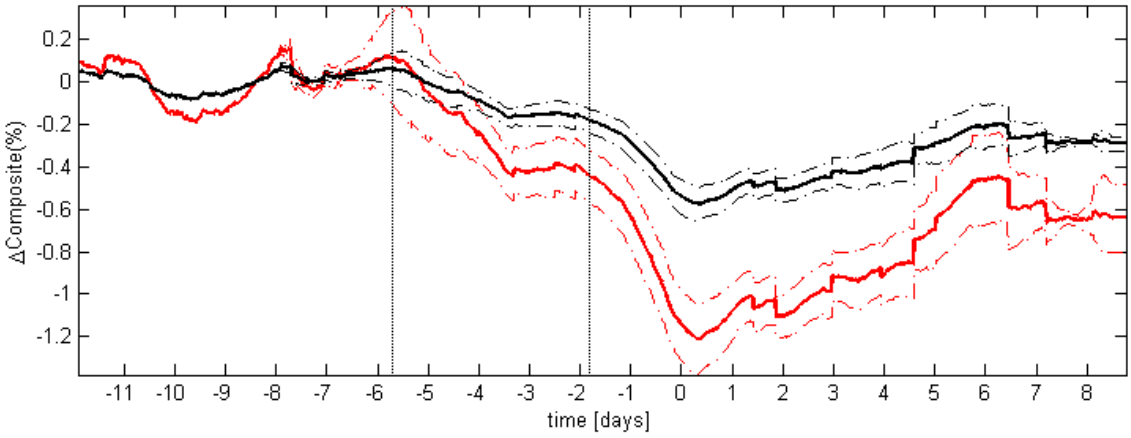


Figure 30 - Average composite measure of variability

In red are displayed the results of the composite; for comparison, in black are displayed the results of the detrended fluctuation analysis area under the curve, after admission condition normalization. The continuous lines represent the average value of the time series across the population, and the dashed lines represent plus or minus the standard error of the mean. The two vertical dotted lines highlight when, on average, the composite variability started to drop. Before averaging, for each of the 14 subjects developing sepsis the time series of either the composite or the detrended fluctuation analysis were aligned to the time of administration of antibiotics (t = 0). The picture shows the higher sensitivity of the composite to sepsis development, respect to the sensitivity of a single HRV measure.

Results

Composite measure of variability

To create a measure tracking the physiological condition of patients undergoing sepsis development, two major signal processing steps are introduced after the computation of 92 variability measures: 1) selection of the informative measures based on nonlinear correlation assessed with a prototype function reflecting decreasing variability, and 2) projection of the selected variability measures onto the first principal component of the population. In the first step, the Spearman's nonlinear correlation between the time course of the measure and the prototype function is computed for every measure from every subject; then, the average population correlation was computed through bootstrap. The thresholding procedure (see Methods), which aims to retain 10% of the measures with the higher correlation, resulted in 11 surviving variability measures (see Table 17 for details).

The survived techniques were projected onto the population first principal component, which on average accounted for 95.6% of the variance of the measures. Those techniques found different relevance inside the first principal component, as showed in Figure 31.

Table 17 - Selected measures of variability

Measure number	Measure name	Short description
1	Standard deviation	Measure the dispersion of the data from its mean value.
2	Coefficient of variation	Ratio between the standard deviation and the mean of the distribution.
3	Power law Y intercept	After the power spectrum of the time series is computed, a line is fitted in the frequency range [10^{-4}, 10^{-2}] Hz. This measure is the value of the intercept of the y-axis of that line. The power spectrum was computed using the Welch's method on the interpolated R-R interval, after spline interpolation at 4 Hz.
4	Detrended fluctuation analysis area under the curve	This measure computes how the variance of the signal change within certain time scales. The area under the curve is the trapezoid integral of the variance-time scales curve.
5	Wavelet area under the curve	Area under the curve of the Wavelet spectral density [202].
6	Shannon entropy	Measure of the degree of complexity of a time series, is based on a weighted sum of the probability of occurrence of a certain.
7	Plotkin-Swamy average energy	The PS energy operator provides a nonlinear estimate of the energy of the signal at a given time. This measure is the average over that energy.
8	Fuzzy entropy	Similarly to sample entropy, fuzzy entropy computes the conditional probability that a pattern seen in an m-dimensional space, could be seen in a (m+1)-dimensional space. The difference is that, to assess whether two points in the phase space are close, a fuzzy membership function is used instead of a Heaviside step function.
9	Correlation dimension Global	Measure of the dimensionality of a time series attractor.
10	Cardiac vagal index	This measure is a combination of the two orthogonal spreads in the Poincaré plot.
11	Largest Lyapunov exponent	Measure quantifying the chaoticity of a system.

For further details about these measures refer to [26]. The specific parameters used to compute the measures are available upon request.

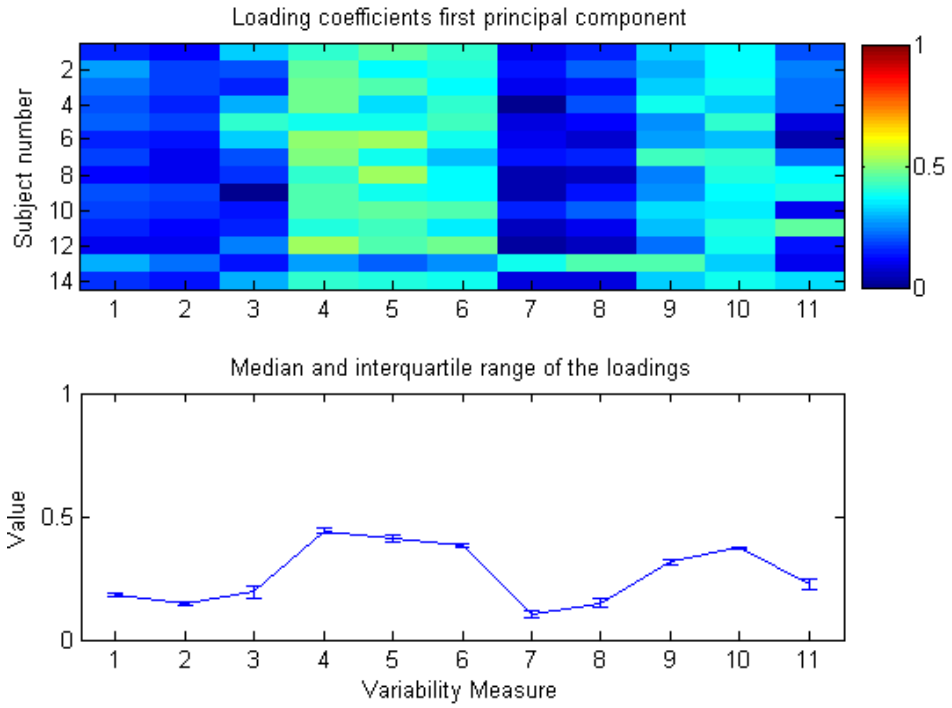


Figure 31 - Population PCA

These panels summarize the model built from the PCA. The list of the eleven variability measures employed in the model, together with their relative number and short description, is reported in Table 17. The upper panel above shows the values of the loading coefficients of the first component for each of the 14 subjects who developed sepsis, and each of the 11 measures which survived the nonlinear correlation selection. Each measure contributed slightly differently to the PCA, depending on the patient. The lower panel represents the same data in the form of median and interquartile error after bootstrapping 1000 times the loading coefficients distributions. Please note that the number identifying the subjects in the figure above (y-axis) do not correspond to the subject numbers reported in [62]; indeed this numbering excludes the subjects who did not develop sepsis.

Looking at the median of the loading coefficients across the population, the area under the curve of the detrended fluctuation analysis (selected measure number 4), the wavelet area under the curve (measure number 5) and the Shannon entropy (measure number 6) were identified as having the highest values. However, no real major contributor to the principal component was found, because all the techniques had close loading coefficients.

The composite measure of variability created through this process presented a clear decline over time for patients who developed sepsis (see Figure 30).

The decline started on average around 140 hours prior the administration of antibiotics, increasing its speed at about 42 hours prior. For comparison we also show the composite for the three subjects who did not develop sepsis in Figure 29.

The composite in that case showed positive values, and was either increasing or oscillating around positive values. The comparison of the composite variability with its single components showed that the projection on the principal component produced changes 2 to 3 times larger than any single measure. For simplicity we reported only the change of the detrendended fluctuation analysis area under the curve (DFA AUC, i.e. measure #4), being the measure with a higher weight in the PCA model, and therefore a major contributor to the composite (see Figure 30).

Detection of sepsis development

To detect sepsis development the summed energy at the very low frequencies (E_s) was extracted from the composite. This feature represents the energy at the frequencies with the same time scale of sepsis development. The detection of sepsis development was achieved by training a logistic regression with two classes of data, one representing early health conditions (E_s of the second day after admission, called E_{s1}), the other representing sepsis at its most advanced stage (E_s from second day to administration of antibiotics,

called E_{s2}). The two classes show considerably different values (Figure 32), making the classification task easier.

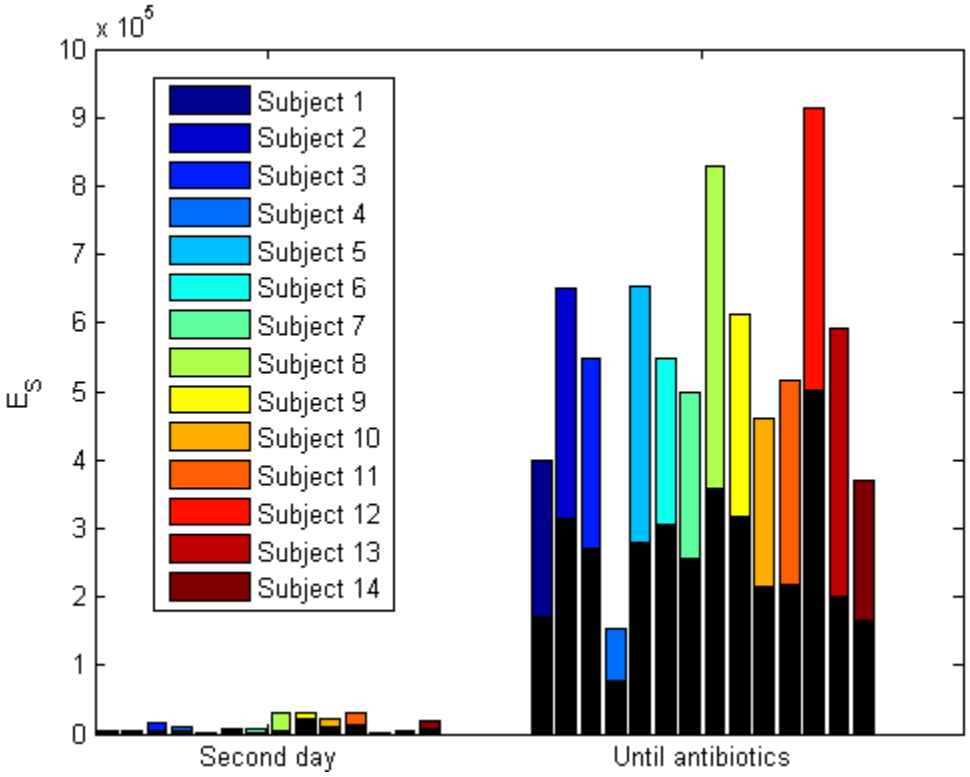


Figure 32 - Energy of the composite measure

The bars represent the energy at the very low frequencies (E_s) of the composite measures for all the subjects during the second day, i.e. E_{s1}, and from the second day to antibiotics administration, i.e. E_{s2}. In colors are the energies extracted from the composite measure of variability, in black are over-imposed the energies extracted from the detrended fluctuation analysis area under the curve after admission condition normalization. The composite showed a larger separation between the two classes, and therefore increased sensitivity to sepsis development. A Wilcoxon signed-rank test showed that the null hypothesis of equal median between E_{s1} and E_{s2} is rejected, for both the composite and the single measure of variability (p-value $\sim 10^{-4}$).

We also computed the energy at the very low frequency for the single measures of variability composing the composite, after admission condition normalization. All of them showed a considerably lower separation between the two classes. Nevertheless, in all cases (composite and single measures) the differences between E_{s1} and E_{s2} allowed

rapid detection of the development of sepsis when tracking over time the probability estimated from the logistic regression (Figure 33).

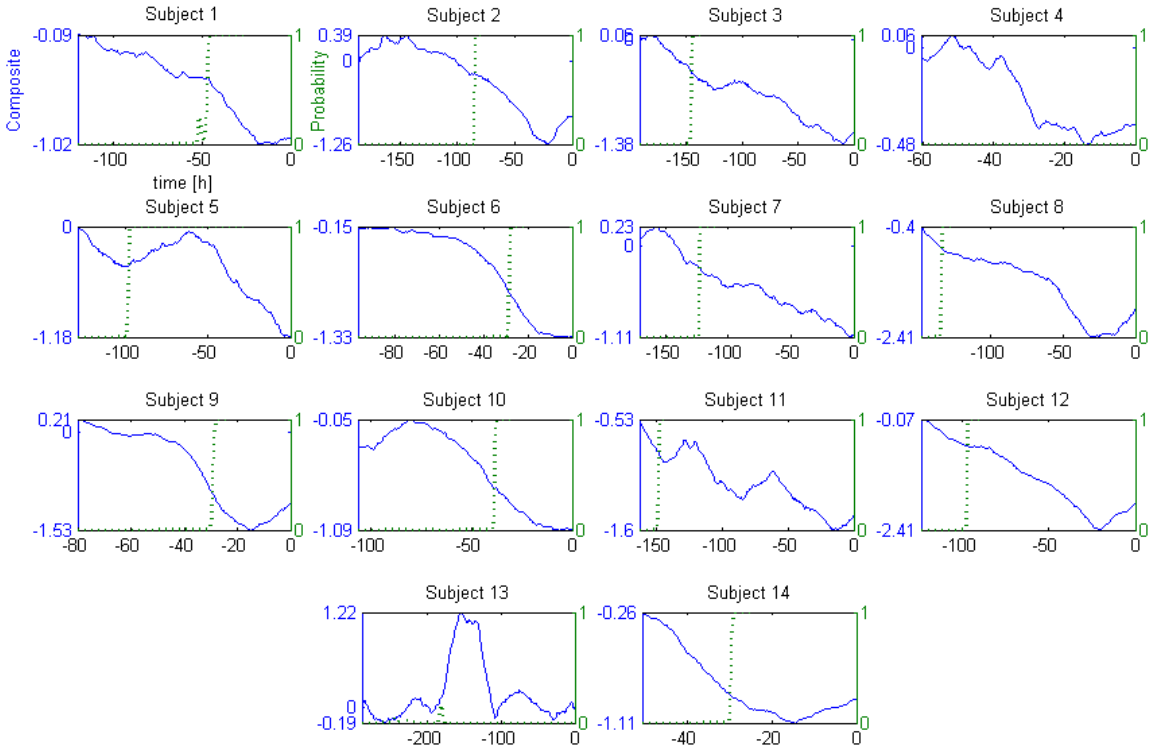


Figure 33 - Probability of development of sepsis
This set of double graphs show the composite measure of variability (blue solid line) and the probability of developing sepsis (green dotted line) at a given time, for each subject. As reported for the plot of subject 1, the x-axis is the time with respect to the administration of antibiotics (t = 0).

This was true for all the subjects except subjects 4 and 13; this result is discussed below. The median time of detection of sepsis development was 60 hours in advance (the values for each subject are reported in Table 18).

Sepsis development was considered "detected" when the probability reached values higher than 0.99. The transition time from probability zero to detection had a median

value of 3.3 hours. The early warning system was run also on the three subjects who did not develop sepsis; as expectable, the estimated probability of developing sepsis remained stable at zero for all of them. The classification performances of the system were: sensitivity 86% (12/14), specificity 100% (3/3), positive predictive value 100% (12/12), and negative predictive value 60% (3/5).

Table 18 - Detection of sepsis development through composite variability

Subject number	Detection time [hours in advance]	Transition time [hours]
1	45.42	7.91
2	84.17	0.41
3	144.2	2.9
4	0	-
5	96.25	3.33
6	27.25	2.33
7	121.7	2.1
8	133.3	1.7
9	27.92	2.5
10	37.5	2.08
11	147.5	2.5
12	96.25	1.25
13	0	-
14	29.17	1.25

Median value of detection 64.7 hours in advance, with a transition time from probability zero to probability higher than 0.99 of 3.3 hours (median value).

Discussion

In this paper a procedure to create a composite measure tracking the change in variability during the development of sepsis was presented. Being based on a dataset of ambulatory outpatients, the procedure makes the major assumption that a patient started being monitored before he actually developed sepsis, or at an early stage of sepsis development. This assumption limits the applicability of the procedure to only a patient population

which starts from a “baseline of health”. However, this specificity is expected to provide improved outcome because taking advantage of a piece of information, i.e. the drop in variability, which would not be present if the patient was already in critical conditions, maybe because of multiple organ dysfunction syndrome (therefore presenting already a low variability).

The creation of the composite measure targeted to sepsis development required two major steps: 1) a selection based on nonlinear correlation with a prototype function representing decreasing variability, which we hypothesize is related to altered clinical physiological state; and 2) the projection on the first principal component of the population of the selected measures, which magnified the information relative to the change in variability. Comparing the trends of the composite measure for both subjects who did (Figure 30) and did not develop sepsis (Figure 29) a distinction appeared, both from the point of view of the trends (the former decreasing, the latter increasing or stable), and of the values of the measure (the former negative, the latter positive). Our system was able to properly identify 15 subjects out of 17, both developing and not developing sepsis. As discussed previously, these results are supported by a growing literature documenting a reduction in HRV correlates with infection and its severity [203]. Other recent examples include the use of short-term heart and respiratory rate variability, combined with other physiological parameters, to predict the risk of severe morbidity due to infection on 138 preterm infants [204], and the use of HRV to predict post-stroke infection [205].

To motivate the creation of the composite measure, for each subject we compared the relative change (i.e. after admission condition normalization) in the composite with the

relative changes of the single measures composing it. We found the composites showed larger changes respect to the single measures. This is not surprising given that the first principal component is a weighted sum of those single changes. The result was confirmed also by the comparison of the sum of the energy at the very low frequencies. This makes the principal component analysis a preferred tool to create a composite measure of variability for this specific application.

To further illustrate the potential clinical utility of the composite measure, we also proposed an early warning system based on the sum of the energy of the composite. We selected the sum of the energy of the composite as the feature of interest after the extraction of a variety of other features, which however did not provide the same results in terms of quality of the prediction of sepsis (see next). Indeed, the chosen system provided in the majority of the cases smooth probability transitions, detecting several hours in advance the development of sepsis. An exception was subject 4, the one with the lowest E_s (Figure 32). Because his E_s was not presented to the logistic regression at the time of creation of the model (we used the leave-one-out cross validation), the decision boundary of the classifier resulted in a value higher than his E_s; this produced a probability of zero. Therefore, this false negative represents an effect of the small sample size. The second exception is subject 13, and is due to the fact that his composite measure had for the majority of the recording positive, and increasing values (Figure 33). In this case a drop in variability did not correspond to physiological deterioration, rather the opposite. Further investigations are needed to explain why this subject expressed a different trend in HRV. Nevertheless, the present study shows how the hypothesis of

direct proportionality between variability and health might be usefully employed for individual patients.

The composite measure showed improved results (i.e. faster detection time and shorter transition time) respect to the measures with a lower weight in the PCA, and results comparable to the ones of the measures with a large weight in the PCA (such as DFA AUC). This is related to the fact that the observed changes in HRV associated to sepsis resulted large enough not to require the increased sensitivity of the composite (i.e. there was a significant separation between the distributions of Figure 32). Nevertheless, this sensitivity would prove useful in detecting sepsis development when more subtle changes in HRV appeared. For instance subjects developing sepsis faster than what we observed with this dataset would present a smaller energy at the very low frequencies, therefore making the increased sensitivity of the composite useful in better detecting sepsis development.

There are several limitations to this study. While its purpose is principally to introduce a method to integrate numerous variability metrics into a single composite measure, the clinical evaluation is limited to a small pilot dataset. Furthermore, during the design we made a few arbitrary choices which require validation. Some were justified by intuition, such as the choice of the frequency range to compute the E_s, identified by inspection of Figure 27. Others were justified based on results across multiple trials. For example, we selected different filter lengths, types of normalization and number of variability measures to include in the principal component model to improve the smoothness of the profiles of the composite measure and of the probability of developing sepsis. A wider

analysis that better motivates these choices would be beneficial not only in further improving the performance, but also in understanding the independent value of each measure of variability. Furthermore, as mentioned, the proposed results are based on a pilot study, involving only few patients. Thus, the usefulness of the proposed approach needs to be proven on larger datasets.

In summary, the composite measure here proposed represents a tool which joins the clinically valuable information of several measures of variability, in a specific way targeted to enhance the sensitivity to sepsis development. This approach addresses a key challenge in variability analysis, which is the reduction of the dimensionality of the analysis, given by the large number of measures currently available [26]. The idea of merging the information from different measures through data reduction techniques could be applied in future studies to merge information streams from multiple organs. This study, for the first time, presents a methodology with which to combine multiple variability metrics in a feasible and potentially clinically useful manner.

5.3 Manuscript 6: Do heart and respiratory rate variability improve prediction of extubation outcomes in critically ill patients?

Published in Critical Care Forum in April 2014.

My contribution: plan and execution of the analysis, writing of the manuscript in coordination with Andrew JE Seely, review and final approval.

Do Heart and Respiratory Rate Variability Improve Prediction of Extubation Outcomes in Critically Ill Patients?

Andrew JE Seely[1]
Andrea Bravi[2]
Christophe Herry[1]
Geoffrey Green[1]
André Longtin[2]
Tim Ramsay[1]
Dean Fergusson[1]
Lauralyn McIntyre[1]
Dalibor Kubelik[1]
Anna Fazekas[1]
Donna E. Maziak[1]
Niall Ferguson[3]
Samuel M. Brown[4]
Sangeeta Mehta[5]
Claudio Martin[6]
Gordon Rubenfeld[7]
Frank J Jacono[8]
Gari Clifford[9]
John Marshall[10]

[1] Ottawa Hospital Research Institute
[2] University of Ottawa
[3] University Hospital Network, University of Toronto
[5] Mt Sinai, University of Toronto
[6] London Health Sciences Center
[7] Sunnybrook Hospital, University of Toronto
[8] University Hospital Case Medical Center, Case Western Reserve University
[9] University of Oxford, Kellogg College
[10] St. Michaels Hospital, University of Toronto

Abstract

Introduction: Prolonged ventilation and failed extubation are associated with increased harm and cost. The added value of heart and respiratory rate variability (HRV and RRV) during spontaneous breathing trials (SBTs) to predict extubation failure remains unknown.

Methods: We enrolled 721 patients in a multi-center (12 sites), prospective, observational study, evaluating clinical estimates of risk of extubation failure, physiologic measures recorded during SBTs, HRV and RRV recorded before and during the last SBT prior to extubation, and extubation outcomes. We excluded 287 patients because of protocol or technical violations, or poor data quality. Measures of variability (97 HRV, 82 RRV) were calculated from electrocardiogram and capnography waveforms followed by automated cleaning and variability analysis using Continuous Individualized Multiorgan Variability Analysis (CIMVA^{TM}) software. Repeated randomized sub-sampling with training,validation, and testing were used to derive and compare predictive models.

Results: Of 434 patients with high quality data, 51(12%) failed extubation. 2 HRV and 8 RRV measures showed statistically significant association with extubation failure ($p<0.0041$, 5% false discovery rate). An ensemble average of five univariate logistic regression models using RRV during SBT, yielding a probability of extubation failure (called WAVE score), demonstrated optimal predictive capacity. With repeated random

sub-sampling and testing, the model showed mean receiver operating characteristic area under the curve (ROC AUC) of 0.69, higher than heart rate (0.51), rapid shallow breathing index – RSBI (0.61) and respiratory rate (0.63). After deriving a WAVE model based on all data, training-set performance demonstrated that the model increased its predictive power when applied to patients conventionally considered high risk: in patients with RSBI>105 or perceived high-risk of failure, the fold increase in risk of extubation failure was 1.5 (CI:1.05,2.04) and 1.2 (CI:0.77,1.78) vs. 3.0 (CI:1.21,5.24) and 3.5 (CI:1.88,5.38) for WAVE Score<0.5 vs. >0.5, respectively (p=0.09; p<0.01).

Conclusions: Altered HRV and RRV (during the SBT prior to extubation) are significantly associated with extubation failure. A predictive model using RRV during the last SBT provided optimal accuracy of prediction in all patients, with improved accuracy when combined with clinical impression or RSBI. This model requires a validation cohort to evaluate accuracy and generalizability.

Number of words in the abstract: 346 (limit 350)

Trial registration: ClinicalTrials.gov NCT01237886

Key words: heart rate variability, respiratory rate variability, extubation failure, predictive modeling, clinical decision support

Introduction

The clinical decision to extubate an intensive care unit (ICU) patient is critical to both the quality and efficiency of care. Early extubation is desirable to decrease the risks of prolonged intubation, including progressive respiratory muscle weakness[206], risk of ventilator associated pneumonia[207], and increased health care expenditures[208]. Conversely, clinicians aim to limit or avoid failed extubation (usually defined as re-intubation within 48 hours of extubation), as it is associated with increased mortality, length of stay, and cost, as well as greater need for long term rehabilitative care[209,210]. Failed extubation can lead to worse outcomes because of complications that occur at the time of re-intubation, especially if performed emergently, including an adverse impact of prolonged intubation, and deterioration prior to re-intubation[211]. The mortality risk associated with failed extubation is variable and dependent on the reason for re-intubation, with airway obstruction, aspiration, or secretions carrying a lower risk than pneumonia or heart failure [212]. Further compounded by projected increasing costs for care of the critically ill [213], there is a need for improved strategies to reducing the duration of mechanical ventilation while simultaneously avoiding failed extubation[214].

Spontaneous breathing trials (SBTs) - short duration trials of reduced ventilatory support to simulate the increased work of breathing after extubation - are widely used to evaluate readiness for extubation [215]. A variety of parameters including respiratory rate (RR), tidal volume (TV), rapid shallow breathing index (RSBI = RR/TV or "Tobin Index"[216]), airway pressure during the first 100 ms of inspiration (P0.1), partial pressure of arterial

oxygen to fraction of inspired oxygen ratio (P/F), maximal inspiratory or expiratory pressure (MIP or MEP), and cough strength have been evaluated as indicators of extubation readiness [216–218]. In the largest multicenter study of this question, factors that independently increased risk of extubation failure included an elevated RSBI during SBT, positive fluid balance and history of pneumonia [218]. Current recommendations for extubation include a 30-120 minute SBT during which multiple physiological parameters are used to assess whether the SBT is a pass, fail or equivocal [219]. However, multiple international studies demonstrate that 10-15% of ICU patients fail extubation and require re-intubation within 48-72 hours, with rates between 25-30% in high risk patients [210,216,220,221].

Complex systems analysis has been increasingly used to characterize biological phenomena. The manifestation of complex systems behavior is evident in the high degree and complexity of variability in the time series of inter-beat intervals (i.e. interval between successive R-peaks), called heart rate variability (HRV), or inter-breath intervals (i.e. interval between successive breaths), called respiratory rate variability (RRV). Numerous methods have been developed to characterize variability mathematically. These methods have been applied in diverse clinical studies, demonstrating that healthy biological systems possess innate and highly complex patterns of variability, and illness is associated with altered variability and reduced complexity [14,26,185,222]. A decrease in variability is indicative of reduced adaptability, reflects a "stressed" system [15,191], and has been described as a marker of outcome in multiple pathological states, for example

sepsis[191]. We and others have hypothesized that cardio-respiratory variability might be used as a marker of the ability of the cardiopulmonary system to tolerate the increased workload associated with both an SBT, and subsequently, extubation. In several single center studies, both HRV [43] and RRV [44,223] during SBTs have been shown to be associated with failed SBTs or extubation failure; however, the added predictive value of variability measures over and above existing methods has not yet been evaluated.

The two goals of our study were: (1) to compare variability in patients who passed and failed extubation using a wide array of HRV and RRV measures, and (2) to investigate the added value of HRV and RRV in the prediction of extubation outcomes, both individually and in combination, as compared to commonly used clinical variables, namely heart rate (HR), respiratory rate (RR), tidal volume (TV), and RSBI.

Methods

The Weaning and Variability Evaluation (WAVE) research study was a prospective, blinded observational multicenter cohort study conducted in 12 centers. Research ethics boards at each site waived consent for enrolment in this strictly observational study. A list of participating sites and dates of activation are provided in Table 19, together with the name of the local ethical bodies approving the study. The study was powered based on preliminary data from a single center pilot (n=60), to estimate the fold increase of extubation failure (respect to average failure rate – i.e. 12% in this population) within a margin of error of 10% or less with two-sided α=0.05.

Patients were considered for enrolment when an SBT was planned in anticipation of extubation. Inclusion criteria were: invasive mechanical ventilation for >48 hours, at least partial reversal of the condition precipitating mechanical ventilation, stabilization of other organ systems, toleration of pressure support ventilation ≤14 cm H_2O ($SpO_2 \geq 90\%$ with $FiO_2 \leq 40\%$ and PEEP ≤ 10 cm H_2O), hemodynamical stability (low - phenylephrine < 50 ug/min; norepinephrine < 5 ug/min; dobutamine < 5 ug/kg/min; milrinone< 0.4 ug/kg/min - or no vasopressors), stable neurological status (no deterioration in Glasgow Coma Score during prior 24 hours and if measured, intracranial pressure (ICP) < 20 mmHg), and intact airway reflexes (cough and gag). Exclusion criteria were**:** order not to re-intubate should the patient fail extubation, anticipated withdrawal of life support, known or suspected severe weakness (myopathy, neuropathy or quadriplegia), tracheostomy, atrial fibrillation, and prior extubation during ICU stay.

Table 19 - Sites' enrolment

City	Hospital	Ethical body approving the study	Enrolling since	# Patients enrolled
Ottawa	General	Ottawa Health Science Network Research Ethics Board	Nov 2009	384
Ottawa	Civic	Ottawa Health Science Network Research Ethics Board	Feb 2010	89
Ottawa	Heart Institute	Ottawa Health Science Network Research Ethics Board	Nov 2009	45
London	Health Sciences Center	The University of Western Ontario Research Ethics Board for Health Sciences Research Involving Human Subjects	Jan 2011	41
Toronto	Mount Sinai	Mount Sinai Hospital Research Ethics Board	Jul 2011	61
Ann Arbor, MI	University of Michigan	University of Michigan Medical School Institutional Review Board	Aug 2011	7
Cleveland, OH	University Hospitals Case Medical Center	The University Hospitals Case Medical Center Institutional Review Board	Oct 2011	42
Lebanon, NH	Dartmouth Hitchcock Medical Center	Dartmouth College and Dartmouth-Hitchcock Medical Center Committee for the Protection of Human Subjects	Nov 2011	4
Vancouver	UBC St. Paul's	University of British Columbia-Providence Health Care Research Ethics Board	Jan 2012	27
Billings, MT	Billings Clinic	Institutional Review Board of Billings	Jan 2012	6
Murray, UT	Intermountain Medical Center	Intermountain Healthcare Urban Central Region Institutional Review Board	Jul 2012	5
Toronto	St. Michael's	St. Michael's Hospital Research Ethics Board	Sep 2012	10
			Total	**721**

Case Report Forms (CRF)

Research teams at each site screened daily to identify study participants, and completed clinical case report forms (CRFs). Respiratory therapists (RTs) performed the SBTs and completed the SBT and Extubation CRFs. The SBT CRF (one per SBT) included ventilator settings (pressure support, PEEP prior to and during the SBT, inspired O_2 fraction, average tidal volume, and minute ventilation), RR, heart rate (HR), blood pressure, O_2 saturation and RSBI at the 2 min, 15 min, 30 min, and end time of the SBT. The Clinical CRF (one per patient) included patient demographics, ICU admission diagnosis, co-morbidities, APACHE II on the day of admission, date and time of extubation, survival status 30 days after ICU admission, need for tracheostomy, and the need and etiology for reintubation (along with reintubation date and time). Immediately prior to extubation and following a decision to extubate, research personnel completed an Extubation CRF. This form recorded the treating team's clinical perception of risk of extubation failure (high> 15%, average 5-15%, low < 5%), as well as factors assessed in considering the patient's readiness for extubation. Failed extubation was defined as re-intubation within 48 hours of extubation.

Signal acquisition and processing

RTs attached CO_2 modules to the bedside monitor and affixed CO_2 tubing to the ventilator circuit at least 30 min prior to SBT. Waveform data collection included ECG

lead II and CO2 data from 30 minutes prior to the SBT until 30 minutes following its conclusion (encompassing the entire SBT).

R-peak to R-peak interval (RRI) time series were extracted from the ECG waveform using a well-known R-peak detection algorithm [224]. Ectopic beats were excluded using beat annotations as well as a threshold-based detection algorithm [225]. Similarly, the time interval between two successive breaths, i.e. inter-breath interval (IBI) time series, was extracted from CO_2 waveforms (125Hz) through standard zero-crossing detection.

Waveform Quality and Variability Analyses

Using Continuous Individualized Multiorgan Variability Analysis (CIMVA^{TM}) software, a set of 97 measures of HRV and 82 measures of RRV (listed in the electronic supplement [173]) was calculated and tracked over time through a windowed analysis of data collected before and during the SBT prior to extubation (i.e. the last SBT). This analysis consisted of 1) taking a window of the RRI/IBI data (5 min for HRV and 15 min for RRV), 2) computing all variability measures for the given window, and 3) repeating the computation on successive windows with a step size of 2.5 minutes for both HRV and RRV. Waveform quality was assessed in an automated fashion for each window; briefly, the quality filtering was based on the morphology of the ECG/CO_2 waveforms, the level of noise/artefacts and the proportion of disconnected/saturated periods [226,227]; this information was used to exclude patients without high quality ECG/CO_2 and RRI/IBI data (see Figure 34 for exclusions due to poor data quality). The outcome of the variability analysis was summarized over two intervals (30 minutes immediately prior to

SBT start and the first 30 minutes of the SBT), by computing the median of each variability time series within these intervals (excluding windows that contained the SBT start time). The change in variability (defined as the median variability during the SBT minus the median variability prior to the SBT (i.e. delta=during–pre), was also computed.

Statistical analysis

In addition to the HRV and RRV measures, study subjects were compared for gender, age, APACHE II score, ICU admission diagnosis, comorbidities and clinician-perceived risk of extubation failure. The chi-square test was used to compare patient proportions, and the Wilcoxon rank-sum test to compare medians, which were reported with 95% confidence intervals (CI). The robust false discovery rate [228] was used for multiple comparison correction (fixing the rate of false positives to 5%).

Predictive Modelling

Because of its simplicity and robustness [229], we utilized an ensemble averaging of univariate logistic regressions. A univariate logistic regression is a model that takes as input a single measure of variability, and provides as output the risk of failing extubation as a number between 0 and 1. The ensemble averaging consists of taking multiple univariate logistic regressions and averaging their outputs, so as to get a more robust estimate of the risk. The output of our predictive model is called the WAVE score and it represents an estimate of the probability of extubation failure, with values closer to zero indicating lower probability and values closer to one indicating higher probability. The

selection of the variables to be included in the model and the unbiased estimation of its performance required the division of the dataset in three sets – training (creation of decision thresholds), validation (for identification of the best performing set of variables) and test (for unbiased performance estimation), with the use of two validation loops (repeated random subsampling), to ensure robustness of the results [162]. The identification of the best performing variables (i.e. feature selection) was based on a greedy optimization on the validation set. In particular, we kept those variables maximizing the sum of two specific measures: the area under the receiver operating characteristic (ROC AUC – used to select measures with high sensitivity and specificity), and the positive predictive value (PPV - used to maximize predictive accuracy of failed extubation). The greedy optimization started from the single univariate logistic regression showing the highest (ROC AUC + PPV) on the validation set. Then, we added to the model, one-by-one, the univariate logistic regression improving the performance on the validation set, until five variables were selected. We imposed to use five variables following the rule of thumb that log(n) variables should be used with a dataset of n samples. Ideally we should have run a cross-validation loop to optimize the number of variables, however that was not suitable because of the low number of patients who failed extubation. The model was then evaluated on the test data for unbiased estimation of its performance. The described process was repeated 500 times to yield a robust estimate of the average performance of the predictive model.

Subsequently, we trained the ensemble average of univariate logistic regressions on all the data and evaluated how it performed on subgroups of patients. Although the results of the model on the same data used to create it (i.e. training set results) are biased, they enable the comparison of the performance across subgroups. In particular, we characterized the risk/fold increase in risk of failing extubation in four subgroups: low vs. high RSBI (threshold of 105 breaths/min/L, consistent with prior studies [216,230]), and low or average vs. high clinician-perceived risk of failure.

Results

We enrolled 721 patients, 60 patients between November 2007 and April 2009 in a run-in pilot, and the remaining 661 between November 2009 and December 2012. See Figure 34 for a flow diagram of the patient selection process. After exclusions, 434 subjects remained (51 (~12%) failed and 383 passed). These 434 subjects constituted the cohort undergoing CIMVA analysis. The 'failed' and 'passed' included and excluded populations were similar, other than the proportion of patients assessed as having low/average/high risk of failing extubation and the values of RSBI and RR at 30 min during the SBT (Table 20).

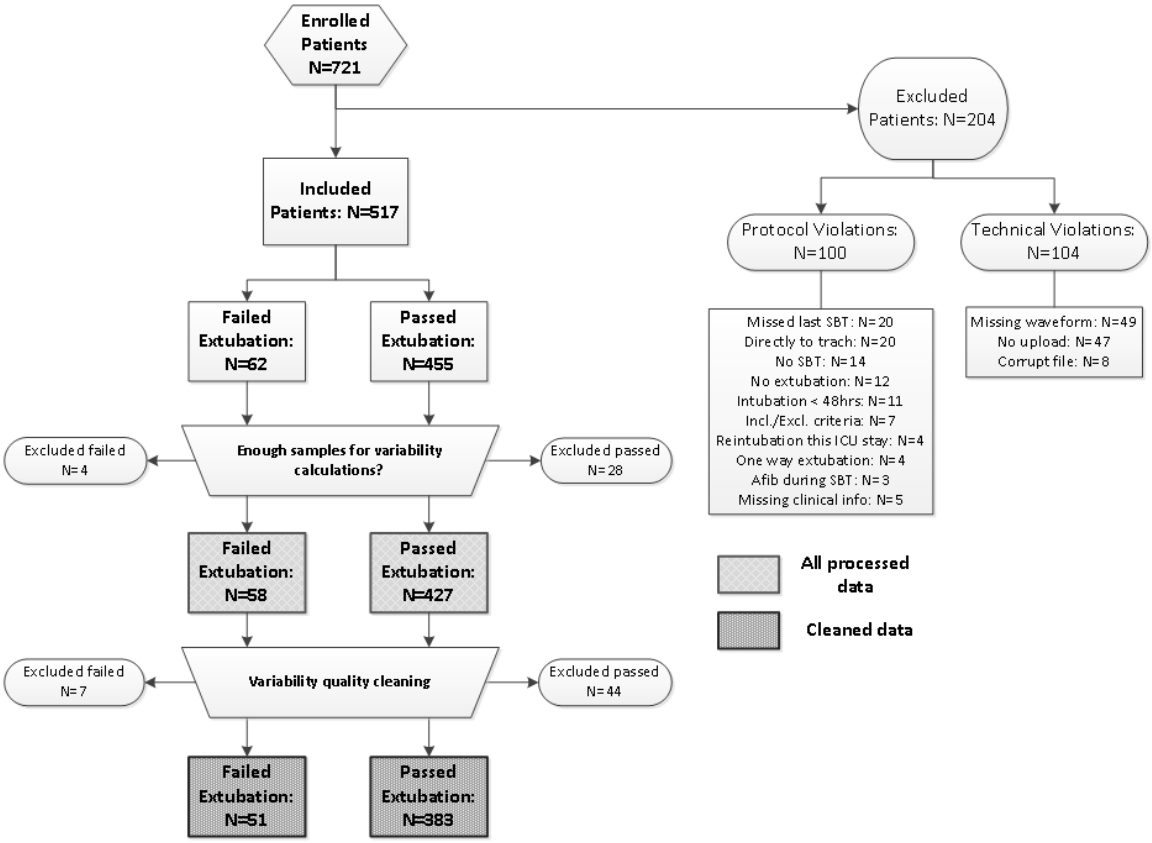


Figure 34 - Flow diagram of selection of patients

Besides standard exclusions due to protocol and technical violations, the diagram shows how the dataset was reduced to ensure proper variability computation. In particular, patients were excluded when 1) having less than two windows of both heart rate and respiratory rate variability to analyze prior and during the spontaneous breathing trial, and 2) variability was extracted from waveforms deemed to be poor quality.

None of the variability measures computed prior to the SBT, nor the difference between variability measures during and prior the SBT, were found to be statistically significant, when adjusted for multiple comparisons using a false discovery rate of 5%. Only the variability measures calculated during the SBT were found significantly associated with extubation failure. As a result, only during SBT variability measures were considered in the subsequent predictive model analysis. Ten measures of variability (2 HRV and 8 RRV) during the SBT were statistically significant ($p<0.0041$ – threshold with 5% false

positives), as summarized in Table 21. A visual representation of the distributions for RR, RSBI and the measure of variability with the lowest p-value (i.e. RRV Recurrence quantification analysis: maximal diagonal line) is provided in Figure 35. Despite the low p-values associated with univariate comparisons, the distributions manifest substantial overlap between the passed and failed categories, highlighting the need for a multivariate predictive model.

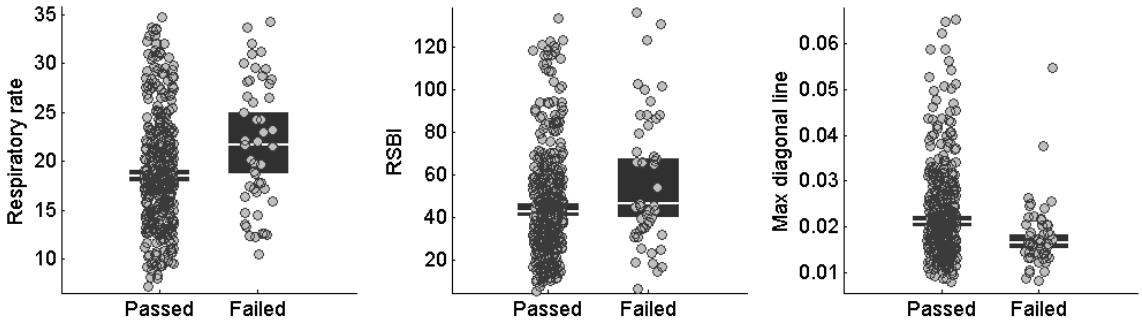


Figure 35 - Distributions of RR, RSBI and variability

This figure shows the distribution of values for Passed and Failed of three different measures (from the left: respiratory rate, rapid shallow breathing index, and RRV recurrence quantification analysis maximal diagonal line). Each grey circle represents a subject. The black box with a white line in between represents the median with its 95% confidence interval.

The comparison of WAVE score (unbiased test set performance results) with logistic regression models based on clinical parameters commonly used to predict extubation outcome is reported in Table 22, showing that RRV variables achieved the highest ROC AUC, demonstrating improved sensitivity (+25%), without substantial changes in positive predictive value (identical) or negative predictive value (+3%).

Table 20 - Patient demographics

	Passed extubation (N=383)^	Failed extubation (N=51)^	p-value*
Gender:			
Males, n (%)	186 (48.6)	29 (56.7)	0.27
Females, n (%)	191 (49.9)	21 (41.2)	0.24
Age (95% CI)	63 (61, 64)	65 (58, 69)	0.86
APACHE score (95% CI)	19 (19, 20)	20 (18, 23)	0.21
Level of sedation[x] (95% CI)	0 (0, 0)	0 (-1, 0)	0.78
ICU Admission Diagnoses			
Cardiovascular, n (%)	112 (25.4)	12 (20.0)	0.40
Respiratory, n (%)	87 (19.7)	18 (30.0)	0.05
Infections, n (%)	62 (14.1)	10 (16.7)	0.54
Gastrointestinal, n (%)	34 (7.7)	3 (5.0)	-
Surgery, n (%)	33 (7.5)	3 (5.0)	-
Head, n (%)	36 (8.2)	1 (1.7)	-
Renal, n (%)	18 (4.1)	2 (3.3)	-
Trauma, n (%)	8 (1.8)	1 (1.7)	-
Overdose, n (%)	9 (2.0)	1 (1.7)	-
Pancreatitis, n (%)	3 (0.7)	1 (1.7)	-
Hepatobiliar, n (%)	5 (1.1)	0 (0.0)	-
Other, n (%)	34 (7.7)	8 (13.3)	0.12
Comorbidities[+]:			
None, n (%)	237 (61.9)	28 (54.9)	0.34
Lung, n (%)	90 (23.5)	15 (29.4)	0.35
Heart, n (%)	81 (21.1)	13 (25.5)	0.48
Both, n (%)	25 (6.5)	5 (9.8)	0.39
Ventilation Settings Pre-SBT:			
PEEP (95% CI) [cm H2O]	10 (8 10)	8 (8 10)	0.90
PS (95% CI) [cm H2O]	10 (10 10)	10 (10 10)	0.14
FiO_2, (95% CI)	30 (30 30)	30 (30 30)	0.04
Ventilation Settings During-SBT:			
PEEP (95% CI) [cm H2O]	5 (5 5)	5 (5 5)	0.46
PS (95% CI) [cm H2O]	5 (5 5)	5 (5 5)	0.01
FiO_2, (95% CI)	30 (30 30)	30 (30 30)	0.11
Perceived risk of failure:			
N/A, n (%)	53 (13.8)	7 (13.7)	0.98
Low, n (%)	117 (30.5)	6 (11.7)	0.005
Average, n (%)	180 (47.0)	26 (51.1)	0.59
High, n (%)	33 (8.7)	12 (23.5)	0.001
Respiratory rate: [breaths/min]			
Pre-SBT (95% CI)	16.0 (16.0, 18.0)	18.0 (15.0, 22.0)	0.09
During-SBT (95% CI)	18.4 (17.9, 19.0)	21.7 (18.7, 25.0)	0.005
RSBI: [breaths/min/L]			
Pre-SBT (95% CI)	34.1 (31.8, 36.4)	40.0 (32.5, 50.0)	0.16
During-SBT (95% CI)	42.7 (39.3, 45.6)	46.6 (40.0, 67.5)	0.005

Low, average and high risk categories were based on clinical impression.

[^]There is a maximum amount of 2% of missing values in each category, due to clinical data not recorded (e.g. Males and Females in the Failed population add up to 50, instead of 51)

* Comparison between Passed and Failed using Wilcoxon rank-sum test to compare the medians, or Chi-square test to compare the proportions (no p-value was provided for those variables with less than 5 samples)

[x] Computed through Richmond Agitation Sedation Scale (RASS) score

[+] Heart comorbidities are coronary artery bypass graft, dilated cardiomyopathy, congestive heart failure, and coronary artery disease; Lung comorbidities are pulmonary fibrosis, COPD, asthma; "None" corresponds to no lung or heart comorbidities.

Table 21 - Statistically significant comparisons of during-SBT variability

Variability Domain	Measure name	Passed (n=383)	Failed (n=51)	p-value
Statistical	HRV Mean of the differences	$1.4\ 10^{-6}$ ($-9.4\ 10^{-7}$, $3.7\ 10^{-6}$)	$-8.4\ 10^{-6}$ ($-1.5\ 10^{-5}$, $-1.9\ 10^{-6}$)	0.00278
Geometric	RRV Recurrence quantification analysis: average diagonal line	0.0057 (0.0054, 0.0060)	0.0044 (0.0038, 0.0053)	0.00011
	RRV Recurrence quantification analysis: maximum diagonal line	0.021 (0.020, 0.022)	0.016 (0.015, 0.018)	0.00004
	RRV Recurrence quantification analysis: maximum vertical line	0.017 (0.016, 0.018)	0.012 (0.011, 0.014)	0.00017
	RRV Recurrence quantification analysis: trapping time	0.0048 (0.0046, 0.0050)	0.0038 (0.0030, 0.0042)	0.00009
Informational	RRV Fano factor distance from a Poisson distribution	-0.12 (-0.12, -0.11)	-0.15 (-0.17, -0.12)	0.00166
Energetic	RRV Hjorth parameters: activity	11.1 (10.4, 11.8)	7.8 (6.0, 10.7)	0.00406
Invariant	HRV Power Law (based on frequency) x intercept	15.8 (14.8, 17.3)	10.0 (4.5, 13.9)	0.00255
	RRV Largest Lyapunov exponent	1.02 (1.00, 1.02)	1.07 (1.03, 1.14)	0.00151
	RRV Power Law (based on histogram) y intercept	-2.17 (-2.21, -2.10)	-2.35 (-2.59, -2.15)	0.00259

*For specific details about each measure, refer to [173].

Table 22 - Prognostic accuracy comparison

Model	ROC AUC	Sensitivity*	Specificity*	PPV*	NPV*	NRI*
Single logistic regression: Heart Rate	0.51	0.5	0.55	0.11	0.87	0.22
Single logistic regression: RSBI	0.61	0.5	0.72	0.18	0.91	0.04
Single logistic regression: Respiratory rate	0.63	0.5	0.66	0.17	0.91	0.04
Ensemble of three univariate logistic regressions: Heart rate, respiratory rate, RSBI	0.62	0.5	0.69	0.17	0.90	0.07
WAVE score	0.69	0.75	0.59	0.18	0.94	-

* Positive test for probability of failure equal or above 0.5
PPV – positive predictive value; NPV – negative predictive value; NRI – net reclassification improvement

We further characterized the WAVE score by assessing its training set performances 1) in the whole population stratified in quartiles, and 2) in association with RSBI and clinical impression of perceived extubation risk, using a decision threshold of 0.5 on the WAVE score. The risk of failure was defined as the number of patients who failed divided by the total number of patients in a given group. The fold increase in risk is the risk divided by the average risk of failure of the dataset (~12%). The training set performances on the

entire dataset (i.e. no data left in the test set) stratified in quartiles are shown in Figure 36, where we see that for higher WAVE scores there is a corresponding higher risk of failing extubation. Similarly using a binary cutoff, we found a fold increase in risk for WAVE score above 0.5 of 1.59 (95% confidence interval – CI: 1.16, 2.02).

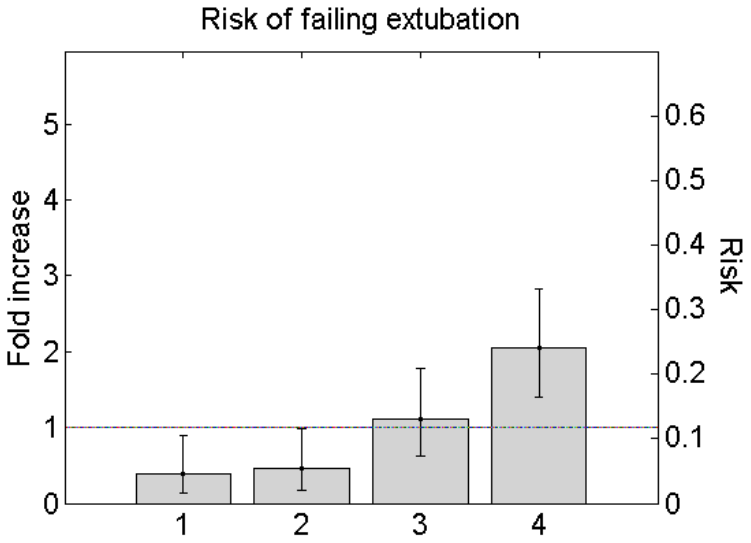


Figure 36 - WAVE score quartile
This figure shows the risk/fold increase in risk of failing extubation associated with each quartile of the population. The risk is defined as the number of patients who failed divided by the total number of patients in a given quartile. The fold increase in risk is the risk divided by the average risk of failure of the dataset (~12%).The total number of patients is 434, therefore each quartile is representative of 108 patients.

In Figure 37, we observed that the higher the RSBI, or the clinician-perceived risk of failure, the stronger the ability of the model to identify extubation failure. In particular, the fold increase in risk for WAVE score above 0.5 moved from 1.5 (CI: 1.05, 2.04) for patients with RSBI<105 to 3.00 (CI: 1.21, 5.24) for patients with RSBI>105. Similarly, the risk increased from 1.21 (CI: 0.77, 1.78) for patients with a low/average perceived risk of failure to 3.54 (CI: 1.88, 5.38) for those having a high perceived risk of failure.

We performed the same comparison on the ROC AUCs. The ROC AUC of the WAVE score on the entire dataset was 0.72. For the subgroups we achieved instead (in order): 0.67 for patients with low/average perceived risk of failure, 0.69 for those with low RSBI, 0.82 for patients with high perceived risk of failure, 0.87 for those with high RSBI ($p<0.01$ and $p=0.09$, respectively).

Discussion

This multicenter observational study demonstrates that in a broadly heterogeneous population of critically ill patients requiring mechanical ventilation, both HRV and RRV during the last SBT prior to extubation are associated with subsequent extubation failure. Two measures of HRV and eight measures of RRV recorded during the last SBT prior to extubation showed statistically significant differences with respect to extubation outcome, and to a greater degree than HR, RR or RSBI.

Using a machine learning analysis with randomized repeated sub-sampling and cross-validation, the average predictive capacity of RRV during the last SBT was superior to all other measures. The use of several measures of RRV to derive the WAVE score showed a higher ROC AUC, outperforming RR, HR and RSBI. Standard choices for the parameters of the repeated random sub-sampling validation (80% training, 10% validation, 10% test)[162] and the decision threshold to compute sensitivity, specificity, negative/positive predictive value (threshold=0.5) were chosen to reduce bias. The training set performances of the WAVE score were used to evaluate sub-group

performances to get a sense of the way the model might be used clinically. The ROC AUC was 0.87 and 0.82 in the high-risk patients, based on RSBI>105 and high perceived risk of extubation failure, respectively.

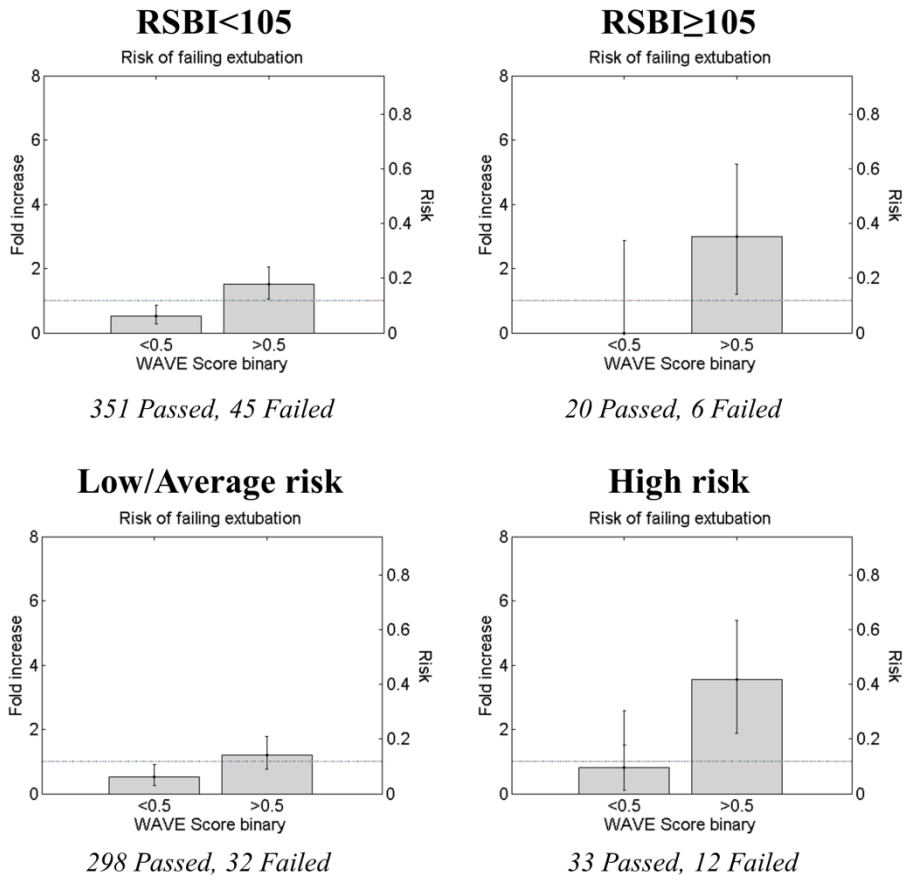


Figure 37 - WAVE score, RSBI and clinical impression

These figures show how the risk/fold increase in risk of failing extubation associated with positive WAVE score (i.e. above 0.5) increases with increasing RSBI during SBT (above), or the clinical impression of the physician at the end of the SBT (below). The risk is defined as the number of patients who failed divided by the total number of patients in a given group (e.g. above 0.5). The fold increase in risk is the risk divided by the average risk of failure of the dataset (~12%). There are 396 patients with low RSBI (45 Failed, 351 Passed), and 26 patients with high RSBI (6 Failed, 20 Passed), while 12 Passed had no RSBI reported. There is no statistically significant difference between the number of Failed and Passed that had no RSBI reported (p-value = 0.2, chi-squared test for proportions). There are 330 patients with low/average risk of failure (32 Failed, 298 Passed), and 45 with high risk of failure (12 Failed, 33 Passed), while 7 Failed and 52 Passed have no perceived risk of failure reported. There is no statistically significant difference between the number of Failed and Passed that had no perceived risk of failure reported (p-value = 0.98, chi-squared test for proportions).

In contrast to the average un-biased model performance, these performances are biased because they are tested on the entire dataset from which the WAVE score was derived (i.e. trained and validated). Taken together, the results demonstrate that RRV during the last SBT outperforms any other measure(s) to predict extubation outcomes, and the sub-group analyses highlight the potential utility and complementarity of the WAVE model for extubation decision making.

The study was inclusive of a highly heterogeneous group of patients (including a wide variation in age, ICU admission diagnosis and comorbidities) in an observational study with absence of strict protocolization of SBT performance (e.g. type, pressure support levels, duration) or extubation decision making; thus, the variation in patients and practice may have diluted the observed signal.

As extubation is associated with an increase in work of breathing [231], and extubation failure is commonly due to the inability of the cardiorespiratory system to tolerate this increased workload, it is not surprising that the commonly utilized quantitative tests for predicting failed extubation [217] (e.g. RSBI) are markers of inability of the cardio-respiratory system to respond to an increased workload. We hypothesized that variability monitoring would improve the ability to detect stress and inability to tolerate the increased workload of breathing associated with an SBT, and subsequently, extubation. Our findings support this hypothesis, and are consistent with prior studies that have demonstrated that HRV and RRV help to predict SBT or extubation outcomes. In 2001,

El-Khatib et al. showed that during the SBT, 13 patients who failed extubation (defined as re-intubation within 24 hours) had lower complexity than 39 patients who were successfully extubated [232]. In 2003 Shen et al. [43] showed that spectral measures of HRV were reduced in 12 patients who failed the weaning trial (either failing extubation or not being ready for extubation). Bien et al. [223] showed in 78 postoperative patients (57 passed weaning, 21 failed) that four measures of RRV, along with RR and RSBI, exhibited significant differences between the two groups. Work by Wysocki et al. in 2006 [44] confirmed the results of Bien in 46 ICU patients (32 passed and 14 failed). Papaioannou et al.[233] published a study on 42 postoperative patients (24 passed weaning, 18 failed), showing that a set of non-linear measures of HRV and RRV provided added value to a predictive model based on RSBI.

The understanding of altered respiratory rate dynamics in association with extubation failure remains an area of active investigation. Of note, the recurrence quantification analysis (RQA) of RRV emerged as highly significantly associated with extubation failure (Table 21). The recurrence plot is a technique that projects a time series (in this case the inter-breath interval time series) in a higher dimensional space, called phase space. In that space the pairwise Euclidean distance between all points is computed, creating a matrix where each row and each column is a point in the phase space, and each element of the matrix is the respective distance. When this distance is smaller than a given threshold, i.e. two points are close in the phase space, a “recurrence” occurs. RQA consists in the study of the number and types of recurrences that appear in a recurrence

plot. Failed extubation appeared to show slightly higher degree of chaotic dynamics, given by the shorter length of diagonal and vertical lines, as compared to Passed extubation. This result is supported by an increase in the largest Lyapunov exponent, a measure of chaoticity of a system. These findings are in keeping with shorter RQA diagonal lines demonstrated during 39 failed SBTs compared to 92 successful SBTs (albeit with no study of extubation) [234]. For more details on these measures, refer to [26].

Given that prior studies utilized visual inspection of waveform data to ensure adequate waveform quality, an important strength of this study was the use of automated quality filters. No visual inspection or hand-selected methodology was utilized to choose waveforms or patients for variability analyses. It is well known that artifact, ectopy and non-stationarity can dramatically alter variability measures [235]. Our quality filters were developed based on published electrocardiogram quality algorithms [226,236], as well as proprietary capnography quality filters trained with separate data sets. We verified *a posteriori* that the predictive performance of the model described in this study was significantly lower when including poor quality waveform data.

There are several important limitations to this study, the most significant being its observational derivational design. As such, there was no strict protocolization regarding the way SBTs were performed, nor regarding decisions about extubation. This was a pragmatic observation of a heterogeneous group of patients being considered for extubation. The majority of patients (75.4%) had a ventilator setting of 5 cm H_2O

pressure support and 5 cm H_2O PEEP during the SBT, which may diminish the variability signal (compared to T-piece SBT), reducing specificity [237] and dampening the observed signal within this study. By limiting our analysis to the last SBT preceding extubation, we have utilized the information available to clinicians making the decision to extubate; however there may be information utilized by physicians in following SBT results from day to day, which were not captured in this study.Just over half of all patients were enrolled in a single center, limiting external generalizability; although no significant differences were observed in SBT ventilator settings or model performance of the single site (ROC AUC 0.72) compared to all others (ROC AUC 0.74), a more even distribution of enrolment in a validation cohort is required. A large number of patients were excluded from the analyses, highlighting the challenge in obtaining waveform data, and the potential for a patient to deviate from expected planned extubation. The exclusions appeared random, occurred throughout the study, and equally from all sites; we did not detect any systematic pattern to the exclusions, either technical or protocol violations, or poor waveform quality. Lastly, although to create the WAVE score we made multiple choices to maximize its generalizability - such as using an ensemble of logistic regressions rather a single multivariate, using stratified cross-validation, choosing a non-optimized number of features (i.e. 5 measures of variability), and using the generic decision threshold of 0.5 - there is no guarantee that we did not overfit the predictive model to our population, particularly because of the small number of patients who failed extubation. External validation of the WAVE score is required.

The question of the incremental value of the WAVE score is critically important and complex. Clinical experience and the data shown in this article highlight that extubation outcome prediction is difficult, and no test in isolation is capable of determining extubation outcomes. It is a fully integrated assessment made by ICU clinicians that determines the optimal timing of extubation, based on the assessment of SBT performance, clinical trajectory, comorbidities that affect the risk of harm of extubation failure, patient/family wishes/goals of care and other potential factors. No clinician evaluates a single score to make this complex decision, and the WAVE score is not intended to be used in isolation as a determinant of extubation outcomes. Nonetheless, we believe, that the WAVE score, when used in conjunction with existing measures (e.g. RSBI) provides optimal prediction of extubation outcomes that will be beneficial to the decision making process. For example, the 75% sensitivity of the WAVE score model is higher than the 50% sensitivity of RSBI that we observed in our study. The augmented ROC AUC in high risk patients suggest the complementary utility of avoiding unnecessary delays in extubation when the WAVE score is less than 0.5 or considering alternatives to extubation if the WAVE score is greater than 0.5 in patients deemed high risk based on traditional measures. The patients conventionally identified, as high risk may well be the ones who benefit most from the additional WAVE test. In general, we showed that the WAVE score can stratify patients in multiple categories of risk (Figure 37), thereby providing clinicians with a more representative picture of the status of a patient. The ultimate aim of this research program is to introduce extubation clinical decision support immediately following SBT completion. We anticipate that combining

(1) a standardized method of performing and assessing an SBT, (2) conventional predictive measures, and (3) a novel score like WAVE, will optimally assist the clinician in the decision to extubate. Following a validation study, the true incremental value of this approach would be assessable in a randomized controlled trial.

Conclusions

The determination of optimal timing for extubation of critically ill patients remains an integrated clinical evaluation and assessment made by a clinician at the bedside, with the full understanding of that patient's goals of care, past medical history, etiology of respiratory failure, clinical course in the ICU, as well as their performance on their last SBT prior to extubation. In the largest multicenter study to evaluate the potential and added value of variability in assisting with assessing extubation readiness, we have found that altered HRV and RRV during the last SBT prior to extubation are significantly associated with extubation failure, and a predictive model derived from RRV during SBT provides added prognostic accuracy in predicting extubation failure when compared to physiological variables used in clinical practice, particularly in high risk patients. This model requires validation in an independent cohort to verify its generalizability, and a randomized trial to assess its clinical utility.

Key Messages

- No single measure drawn from SBT performance is capable of accurately predicting extubation outcomes; extubation outcome prediction is challenging.
- Altered HRV and RRV during the last SBT prior to extubation are associated with subsequent extubation failure.
- A multivariate predictive model based on RRV during the last SBT offers improved predictive accuracy of extubation outcomes compared to physiological variables commonly used in clinical practice, particularly in patients deemed high risk of failure based on traditional measures.
- A multicenter validation study is merited and necessary to evaluate the accuracy of the derived predictive model.

Abbreviation list

AUC – Area Under the Curve
CI – Confidence Interval
CIMVA – Continuous Individualized Multiorgan Variability Analysis
COPD – Chronic Obstructive Pulmonary Disease
CRF – Case Report Form
HR – Heart Rate
HRV – Heart Rate Variability
ICU – Intensive Care Unit
MEP – Maximal Expiratory Pressure
MIP – Maximal Inspiratory Pressure
NPV – Negative Predictive Value
PEEP – Positive End-Expiratory Pressure
PPV – Positive Predictive Value
ROC – Receiver Operating Characteristic
RR – Respiratory Rate
RRV– Respiratory Rate Variability
RSBI – Rapid Shallow Breathing Index
SBT – Spontaneous Breathing Trial
TV – Tidal Volume
WAVE – Weaning And Variability Evaluation

Competing Interests

Andrew Seely is Founder and Chief Science Officer of Therapeutic Monitoring Systems (TMS); TMS aims to commercialize patent-protected applications of multiorgan variability monitoring to provide variability-directed clinical decision support at the bedside to improve care for patients at risk for or with existing critical illness. Andrew Seely holds a patent jointly with co-authors Andrea Bravi and André Longtin on composite measures of variability. Geoffrey Green is Product Manager for TMS. John Marshall and Gari Clifford are on the Scientific Advisor Board of TMS. Other authors have no relevant conflict of interest to disclose.

Authors' contributions

AS conceived of the study, and led the grant submissions, study design, recruitment, analyses and writing. AB led the statistical and variability analyses, assisted by CH, GG, AL, TR, GR and DF. DK and AB initiated the draft of the manuscript. DM, TR, DF, LM, GR, GC and NF all provided input into the design, analyses, along with manuscript review and editing. JM, SM, FJ, SB and CM all participated in patient enrolment, as well as providing input regarding design and analyses. AF acted as multisite coordinator for the study. All the authors were involved with the revision of the manuscript and provided final approval.

Acknowledgments

The study was funded by peer-reviewed grants from The Ottawa Hospital Academic Medical Organization (one-year support obtained in 2009) and the Canadian Institutes of Health Research (three-year operating grant obtained in 2010). Both funding sources provided peer-reviewed feedback regarding the design and planned analyses prior to implementation of the study. The following are acknowledged for their invaluable contributions to the study: Daphne Townsend, Peter Dodek, Jon Hooper, Damon Scales, Peter Wilkes, Ryan Zarychanski, Robert Merchant, Athos Rassias, James Blum, and the WAVE Investigators for the Canadian Critical Care Trials Group.

Chapter 6

6 General Discussion

"*How can we bring variability analysis to the bedside?*"

With this question at the center of the thesis we focused on three specific areas of research: 1) methodology - the technical challenges involved with variability analysis, 2) interpretation - the physiological meaning of variability, 3) problem-specific clinical applications.

6.1.1 Methodology

In this thesis we presented a broad overview of multiple measures of variability used in clinical applications, providing their mathematical description, interpretation and references, as well as a novel classification into domains. We saw that this analytical power is helpful in maximizing the chances of identifying clinically relevant information, however coming at the price of increased complexity of the analyses and their interpretation (Chapter 2). It is therefore required to converge the relevant information into a unique index, which we called composite measure, to be easily used by clinicians to take decisions. To do so, we presented a customizable signal processing scheme with emphasis on normalization, feature selection and integration. Each of those three steps can be tuned depending on the prior knowledge on the specific clinical problem, the type of information deemed of interest, and the foreseen application of the composite measure as a decision tool (Chapter 3).

Extraction of measures of variability and integration into a composite measure are two essential steps required to create a decision support system based on variability, however there are several technical details involved that were not a focus of this thesis. We discussed in Chapter 2.1 that the identification of the values for the parameters involved with the extraction of measures of variability is a major challenge. To date the most common approach is to use "optimal parameters" from other studies in the literature. The optimality of a given parameter is usually addressed by trying to maximize the utility of a measure in a given clinical application. The problem with this approach is that a parameter which proved to be optimal to differentiate between, say, congestive heart failure patients and controls, may not be optimal when trying to monitor sepsis development. Potentially, we should derive the "optimal parameter" for any dataset under study, with the challenge that parameter tuning requires the use of a subset of the dataset which cannot be used for any other purpose, thereby making this impossible for small sample size studies. Given the right amount of data, both in terms of size and heterogeneity, a possible approach could be to identify those values of the parameters that produce a global effect, such as for instance the ordering of the data in a monotonic way with the associated degree of impairment. This means a measure would have low values when a patient is more likely to be impaired, and high values when he is more likely to be healthy, thereby making the measure more useful from a clinical perspective. Although that is commonly thought to be the case (i.e. low variability is associated with illness – see Chapter 1.4), smaller studies which were not presented in this thesis showed that this

applies to measures such as standard deviation and mean rate, but does not to the majority of the other measures.

Another problem which was not addressed is that of data quality. To make sure the decision support is reliable, there is the need to know that the underlying data used to create that support is of "enough" quality. Data quality can be extracted at the waveform level (e.g. noise in an electrocardiogram), event level (e.g. number of ectopies in a R-R interval time series) and measure level (e.g. the goodness of fit of a regression line). Those quality measures are themselves domains of research, and offer the same challenges described before, i.e. parameter identification and proper integration into a unique quality score to be easily used to take decisions about the quality of a result.

6.1.2 Interpretation

The physiological interpretation of variability is of primary importance for two main reasons. First, clinicians are more inclined in using tools they understand. Second, the association to physiological phenomena of given measures of variability would enable an easier selection to target specific clinical application. To address this question on heart rate variability we used both synthetic and real patient data (Chapter 4), drawing the following conclusions: 1) the stereotypical association [HRV ↔ Sympathovagal modulation] is too simplistic because the coupling between physiological parameters involved in cardiovascular regulation enables HRV to reflect changes outside of the sympathetic/parasympathetic system; 2) this coupling is the reason for correlation among physiological parameters, and consequently, explains the correlation across seemingly

independent measures of variability; 3) each single measure of variability is therefore a nonspecific sensor of the body; 4) although each measure senses multiple physiological parameters, real patient data comparing changes during sepsis and exercise showed that single measures of variability behave similarly for the two stressors; given that the two stressors have in common only the base effect they produce on straight cardiovascular regulation (e.g. increased cardiac output), to get insights on more hidden physiological parameters it is necessary to perform multivariate analyses; 5) multivariate analysis on the synthetic dataset confirmed this hypothesis, enabling a better tracking of the changes over time in the physiological parameter of interest.

This research highlighted the importance of performing multivariate analysis to get a better understanding of the physiology underlying variability. A future direction could be the identification of the combination of measures leading to the monitoring of specific physiological parameters of interest. To date this is a particularly challenging task because physiologically-based models are not good enough in quantitatively reproducing real data, while real data has artifacts, non-stationarities and confounding factors (e.g. age, gender) further complicating the analyses. A stream of research which seems particularly promising involves bridging the gap between dynamics at the cellular level and systemic level through multi-scale modeling[17]. This has been done in particular to target systemic infection[238–240], linking neuroendocrine-immune interactions with HRV.

6.1.3 Clinical applications

Sepsis development monitoring and extubation outcome prediction were the two clinical applications of variability presented in this thesis (Chapter 5). We showed that HRV can be used to predict about 60 hours in advance the onset of sepsis, and that RRV offers added value to extubation management with respect to commonly used clinical variables, potentially reducing by 9% the extubation failure rate. These results indicate that, when applied at the bedside, these tools could translate in lower lengths of stay, reduced costs and mortality. During the process of deriving these predictive models we identified two main challenges to bring variability at the bedside: 1) definition of normality, 2) sample size calculation.

Clinicians are used to dichotomizing the value of physiological parameters into either normal or abnormal. The normal range of a clinical parameter is known from multiple studies across heterogeneous populations, and provides clinicians with a reference to judge the conditions of a patient. A primary problem with variability is that those ranges are unknown and are hard to define because there is no standard way to perform variability analysis. Furthermore, the length of the data to use (e.g. 5 min, 30 min, 1 hour), how to exclude artifacts and nonstationarities, the values of the parameters used to compute a measure of variability (see Chapter 6.1.1), are all factors affecting a variability measure, altering its range of values. This highlights the relevance of standardizing variability analysis. This standardization would be clinically useful, and would enable a better comparison of results across studies. Given the high number of degrees of freedom

involved with variability analysis, an extensive dataset inclusive of large pools of data from heterogeneous populations will be required to find the optimal standardization strategy.

The second problem we identified is the one of sample size calculation. The standard approach to assess sample size is to plan a statistical analysis, e.g. comparing the median heart rate in two clinically distinct population, and run a power analysis, which is a well-known tool used by statisticians to identify the number of patients that minimizes the chances of type II errors (i.e. failure to reject a false null hypothesis). The problem with it is that power analyses are based on the assumption that the final goal is to run a statistical test, which is a completely different task from creating a model and assessing its predictive power. This is the problem we faced when we found the size of the population used to predict extubation outcome (Chapter 5.3) – i.e. 434 subjects, with 51 who failed extubation – where, despite the relatively large number of subjects, we were underpowered to properly optimize the predictive model. Indeed, the optimization of any parameter involved in a model requires the use of a subset of the data that can be used only for that purpose. In that case we had to divide the dataset into three sets, training (i.e. identification of the threshold), validation (for feature selection) and test (for unbiased estimation of the predictive performance). The additional optimization of one parameter, e.g. the decision threshold which was arbitrarily set to 0.5, would have required the division of the dataset in four sets.

To date there are no ways to estimate the proper sample size to create predictive models. A possibility is to run an adaptive clinical trial, where the sample size is not fixed and Bayesian statistics can be used to assess whether there is an improvement in the predictive performance with an increase in the sample size. If that is the case, it is worth it to keep acquiring subjects until the required performance is obtained. Otherwise, the study can be stopped because not leading to the anticipated results. The two main problems with this approach are that granting agencies have still their mindset tuned to purely statistical studies, and therefore "do not trust" a study that has no fixed sample size, and that there is the objective limitation that is tough to decide the amount of funds to provide if the number of patients is not fixed. Given the proven efficacy of machine learning tools in addressing industrial problems[241], there is the need for additional ways to properly assess sample size.

6.1.4 Contribution

The present thesis has addressed the technology, physiologic understanding and clinical utility of variability. The first contribution (Chapter 2.2) focused on the extraction and categorization of about 100 measures of variability with the purpose to extract all possible pieces of information from the signals under study. Although this practice maximizes the chances of finding relevant information, the challenge becomes dealing with overabundance of information, often correlated and irrelevant. The second contribution (Chapter 3.2) addresses this challenge by providing a procedure to systematically create composite measures of variability addressing specific clinical

problems. The third contribution (Chapter 4.3) shifted the focus on the physiological understanding of variability, comparing changes in variability due to a physiologic stressor (exercise) with those due to a pathologic stressor (sepsis). The results highlighted that single measures of variability are unable to differentiate between the two stressors, although differences were found with a multivariate analysis. This result was confirmed by the fourth contribution (Chapter 4.2). Using a physiologically-based model, the article showed that single measures of variability are mostly independent with each other, as long as there is no change in the parameters of the model. Conversely, a change in the model produced changes in multiple measures of variability, creating correlation. Lastly, the article showed that by combining measures of variability into a unique index in an unsupervised way it was possible to track the changes in the parameters of the model better than what possible using single measures alone. The fifth contribution (Chapter 5.2) addressed the problem of sepsis development, by introducing a composite measure of variability able to faster detect – about 60 hours in advance on average – the development of sepsis. The sixth contribution (Chapter 5.3) addressed extubation management, introducing a decision support score providing better predictive accuracy than any of the clinical measures commonly used to predict extubation outcome.

6.1.5 Conclusion

This thesis represents a small step towards the ultimate goal of any medical research: improving care. We based our work on variability analysis, providing insights on its physiological meaning, how to improve its quantification, and examples of clinical

applications. Several challenges were highlighted, opening opportunities for future research. There is much to do, yet the reward is great. The future is bright.

References

1. Anatomy and Physiology | Primal Online Learning | Anatomy. at <http://www.primalonlinelearning.com/>
2. Seidel, H. & Herzel, H. Modelling heart rate variability due to respiration and baroreflex. *Springer Series in Synergetics* **65,** 205–205 (1995).
3. Stefanovska, A. Physics of the human cardiovascular system. *Contemporary Physics* **40,** 31–55 (1999).
4. Heldt, T., Shim, E. B., Kamm, R. D. & Mark, R. G. Computational modeling of cardiovascular response to orthostatic stress. *J Appl Physiol* **92,** 1239–1254 (2002).
5. Yasuma, F. & Hayano, J.-I. Respiratory sinus arrhythmia: why does the heartbeat synchronize with respiratory rhythm? *Chest* **125,** 683–690 (2004).
6. Opie, L. H. *Heart physiology: from cell to circulation*. (Wolters Kluwer Health, 2004).
7. Katz, A. *Physiology of the Heart*. (Wolters Kluwer Health, 2011).
8. Squire, L. *et al.* in *Fundamental Neuroscience* (Academic Press, 2012).
9. Wolfram, S. Complex systems theory. in *SANTA FE INSTITUTE STUDIES IN THE SCIENCES OF COMPLEXITY-PROCEEDINGS VOLUME-* **1,** 183–183 (1993).
10. Abarbanel, H. D. I. *Analysis of observed chaotic data. 1996*. (Springer, New York).
11. Manuel Varela, Raul Ruiz-Esteban & Maria Jose Mestre De Juan. Chaos, Fractals, and Our Concept of Disease. *Perspectives in Biology and Medicine* **53,** 584–595 (2010).
12. Goldberger, A. L. Non-linear dynamics for clinicians: chaos theory, fractals, and complexity at the bedside. *The Lancet* **347,** 1312–1314 (1996).
13. Seely, A. J. & Christou, N. V. Multiple organ dysfunction syndrome: exploring the paradigm of complex nonlinear systems. *Crit. Care Med* **28,** 2193–2200 (2000).
14. Goldberger, A. L. Heartbeats, hormones, and health: is variability the spice of life? *Am. J. Respir. Crit. Care Med.* **163,** 1289–1290 (2001).
15. Seely, A. J. E. & Macklem, P. T. Complex systems and the technology of variability analysis. *Crit Care* **8,** R367–384 (2004).
16. Cerutti, S., Goldberger, A. L. & Yamamoto, Y. Recent Advances in Heart Rate Variability Signal Processing and Interpretation. *IEEE Transactions on Biomedical Engineering* **53,** 1–3 (2006).
17. Cerutti, S., Hoyer, D. & Voss, A. Multiscale, multiorgan and multivariate complexity analyses of cardiovascular regulation. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **367,** 1337 –1358 (2009).
18. Loscalzo, J., Kohane, I. & Barabasi, A. L. Human disease classification in the postgenomic era: a complex systems approach to human pathobiology. *Molecular systems biology* **3,** (2007).

19. Barabasi, A.-L., Gulbahce, N. & Loscalzo, J. Network medicine: a network-based approach to human disease. *Nat Rev Genet* **12,** 56–68 (2011).
20. Buchman, T. G. Nonlinear dynamics, complex systems, and the pathobiology of critical illness. *Curr Opin Crit Care* **10,** 378–382 (2004).
21. Goldberger, A. L., Rigney, D. & West, B. J. Chaos and fractals in human physiology. *Scientific American* **262,** 42–49 (1990).
22. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Circulation* **93,** 1043–1065 (1996).
23. Kurths, J. *et al.* Quantitative analysis of heart rate variability. *Chaos* **5,** 88 (1995).
24. Rajendra Acharya, U., Paul Joseph, K., Kannathal, N., Lim, C. M. & Suri, J. S. Heart rate variability: a review. *Med Bio Eng Comput* **44,** 1031–1051 (2006).
25. Gang, Y. & Malik, M. Heart Rate Variability: Measurements and Risk Stratification. *Electrical Diseases of the Heart* 365–378 (2008).
26. Bravi, A., Longtin, A. & Seely, A. J. E. Review and classification of variability analysis techniques with clinical applications. *Biomedical Engineering Online* **10,** (2011).
27. Ph.D, P. R. N. in *Complex Systems and Computational Biology Approaches to Acute Inflammation* (Vodovotz, Y. & An, G.) 51–77 (Springer New York, 2013). at <http://link.springer.com.proxy.bib.uottawa.ca/chapter/10.1007/978-1-4614-8008-2_4>
28. Lombardi, F. Clinical Implications of Present Physiological Understanding of HRV Components. *Card Electrophysiol Rev* **6,** 245–249 (2002).
29. Tan, C. O. Heart rate variability: are there complex patterns? *Frontiers in Physiology* **4,** 165 (2013).
30. Godin, P. J. & Buchman, T. G. Uncoupling of biological oscillators: a complementary hypothesis concerning the pathogenesis of multiple organ dysfunction syndrome. *Crit. Care Med* **24,** 1107–1116 (1996).
31. Morris, J. A. *et al.* Reduced Heart Rate Variability: An Indicator of Cardiac Uncoupling and Diminished Physiologic Reserve in 1,425 Trauma Patients: *The Journal of Trauma: Injury, Infection, and Critical Care* **60,** 1165–1174 (2006).
32. Seely, A. J. E. & Macklem, P. Fractal variability: An emergent property of complex dissipative systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **22,** 013108–013108–7 (2012).
33. Csete, M. E. & Doyle, J. C. Reverse Engineering of Biological Complexity. *Science* **295,** 1664–1669 (2002).
34. Doyle, J. C. & Csete, M. Architecture, constraints, and behavior. *Proceedings of the National Academy of Sciences* **108,** 15624–15630 (2011).
35. Baselli, G. *et al.* Spectral and cross-spectral analysis of heart rate and arterial blood pressure variability signals. *Computers and Biomedical Research* **19,** 520–534 (1986).

36. Su, D.-F. & Miao, C.-Y. Blood Pressure Variability And Organ Damage. *Clinical and Experimental Pharmacology and Physiology* **28,** 709–715 (2001).
37. Sega, R. *et al.* Blood Pressure Variability and Organ Damage in a General Population Results from the PAMELA Study. *Hypertension* **39,** 710–714 (2002).
38. Pierdomenico, S. D. *et al.* Prognostic value of different indices of blood pressure variability in hypertensive patients. *Am. J. Hypertens.* **22,** 842–847 (2009).
39. Mancia, G. Prognostic Value of Long-Term Blood Pressure Variability The Evidence Is Growing. *Hypertension* **57,** 141–143 (2011).
40. Hata, Y. *et al.* Office blood pressure variability as a predictor of acute myocardial infarction in elderly patients receiving antihypertensive therapy. *Journal of human hypertension* **16,** 141 (2002).
41. Rothwell, P. M. *et al.* Prognostic significance of visit-to-visit variability, maximum systolic blood pressure, and episodic hypertension. *The Lancet* **375,** 895–905 (2010).
42. Muntner, P. *et al.* The Relationship Between Visit-to-Visit Variability in Systolic Blood Pressure and All-Cause Mortality in the General Population Findings From NHANES III, 1988 to 1994. *Hypertension* **57,** 160–166 (2011).
43. Shen, H.-N. *et al.* Changes of heart rate variability during ventilator weaning. *Chest* **123,** 1222–1228 (2003).
44. Wysocki, M. *et al.* Reduced breathing variability as a predictor of unsuccessful patient separation from mechanical ventilation*. *Critical Care Medicine* **34,** 2076–2083 (2006).
45. White, C. E. *et al.* Lower interbreath interval complexity is associated with extubation failure in mechanically ventilated patients during spontaneous breathing trials. *The Journal of Trauma and Acute Care Surgery* **68,** 1310–1316 (2010).
46. Bradley, B. D., Green, G., Ramsay, T. & Seely, A. J. Impact of sedation and organ failure on continuous heart and respiratory rate variability monitoring in critically ill patients: A pilot study*. *Critical care medicine* **41,** 433–444 (2013).
47. Saria, S., Rajani, A. K., Gould, J., Koller, D. L. & Penn, A. A. Integration of Early Physiological Responses Predicts Later Illness Severity in Preterm Infants. *Sci Transl Med* **2,** 48ra65 (2010).
48. Kelly, G. Body temperature variability (Part 1): a review of the history of body temperature and its variability due to site selection, biological rhythms, fitness, and aging. *Altern Med Rev* **11,** 278–293 (2006).
49. Kelly, G. S. Body temperature variability (Part 2): masking influences of body temperature variability and a review of body temperature variability in disease. *Altern Med Rev* **12,** 49–62 (2007).
50. Varela, M., Jimenez, L. & Fariña, R. Complexity analysis of the temperature curve: new information from body temperature. *Eur J Appl Physiol* **89,** 230–237 (2003).
51. Varela, M. *et al.* Temperature curve complexity predicts survival in critically ill patients. *American journal of respiratory and critical care medicine* **174,** 290–298 (2006).

52. Cuesta, D. *et al.* Predicting survival in critical patients by use of body temperature regularity measurement based on approximate entropy. *Medical and Biological Engineering and Computing* **45,** 671–678 (2007).
53. Papaioannou, V. E., Chouvarda, I. G., Maglaveras, N. K. & Pneumatikos, I. A. Temperature variability analysis using wavelets and multiscale entropy in patients with systemic inflammatory response syndrome, sepsis and septic shock. *Critical Care* **16,** R51 (2012).
54. Tarassenko, L. BIOSIGN^{TM} : multi-parameter monitoring for early warning of patient deterioration. in **2005,** 71–76 (IEE, 2005).
55. Tarassenko, L., Hann, A. & Young, D. Integrated monitoring and analysis for early warning of patient deterioration. *Br. J. Anaesth.* **97,** 64–68 (2006).
56. Watkinson, P. J. *et al.* A randomised controlled trial of the effect of continuous electronic physiological monitoring on the adverse event rate in high risk medical and surgical patients. *Anaesthesia* **61,** 1031–1039 (2006).
57. Fairchild, K. D. & Aschner, J. L. HeRO monitoring to reduce mortality in NICU patients. *Research and Reports in Neonatology* **2,** 65–76 (2012).
58. Moorman, J. R. *et al.* Mortality Reduction by Heart Rate Characteristic Monitoring in Very Low Birth Weight Neonates: A Randomized Trial. *The Journal of Pediatrics* **159,** 900–906.e1 (2011).
59. Glass, L. Introduction to Controversial Topics in Nonlinear Science: Is the Normal Heart Rate Chaotic? *Chaos* **19,** 028501 (2009).
60. Janicke, H., Bottinger, M., Mikolajewicz, U. & Scheuermann, G. Visual Exploration of Climate Variability Changes Using Wavelet Analysis. *Visualization and Computer Graphics, IEEE Transactions on* **15,** 1375–1382 (2009).
61. Lazarou, G. Y., Baca, J., Frost, V. S. & Evans, J. B. Describing network traffic using the index of variability. *IEEE/ACM Trans. Netw.* **17,** 1672–1683 (2009).
62. Ahmad, S. *et al.* Continuous multi-parameter heart rate variability analysis heralds onset of sepsis in adults. *PLoS ONE* **4,** e6642 (2009).
63. Voss, A. *et al.* The application of methods of non-linear dynamics for the improved and predictive recognition of patients threatened by sudden cardiac death. *Cardiovascular Research* **31,** 419 –433 (1996).
64. Schmidt, G. *et al.* Heart-rate turbulence after ventricular premature beats as a predictor of mortality after acute myocardial infarction. *The Lancet* **353,** 1390–1396 (1999).
65. Gil, E., Bailón, R., Vergara, J. M. & Laguna, P. PTT Variability for Discrimination of Sleep Apnea Related Decreases in the Amplitude Fluctuations of PPG Signal in Children. *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING* **57,** 1079–1088 (2010).
66. De Chazal, P., Heneghan, C. & McNicholas, W. T. Multimodal detection of sleep apnoea using electrocardiogram and oximetry signals. *Philos Transact A Math Phys Eng Sci* **367,** 369–389 (2009).

67. Lin, C. L., Wang, J. S. & Chung, P. C. Mining Physiological Conditions from Heart Rate Variability Analysis. *IEEE Computational Intelligence Magazine* **5,** 50–58 (2010).
68. Eckberg, D. L. Sympathovagal Balance : A Critical Appraisal. *Circulation* **96,** 3224–3232 (1997).
69. Singh, R. B. *et al.* Circadian heart rate and blood pressure variability considered for research and patient care. *International Journal of Cardiology* **87,** 9–28 (2003).
70. Bilan, A., Witczak, A., Palusinski, R., Myslinski, W. & Hanzlik, J. Circadian rhythm of spectral indices of heart rate variability in healthy subjects. *Journal of Electrocardiology* **38,** 239–243 (2005).
71. Goya-Esteban, R. *et al.* Heart Rate Variability on 7-Day Holter Monitoring Using a Bootstrap Rhythmometric Procedure. *Biomedical Engineering, IEEE Transactions on* **57,** 1366 – 1376 (2010).
72. Tarvainen, M. P., Niskanen, J. P., Lipponen, J. A., Ranta-aho, P. O. & Karjalainen, P. A. Kubios HRV—A Software for Advanced Heart Rate Variability Analysis. in *4th European Conference of the International Federation for Medical and Biological Engineering* 1022–1025 (2009).
73. Lewicke, A., Sazonov, E., Corwin, M. J., Neuman, M. & Schuckers, S. Sleep versus wake classification from heart rate variability using computational intelligence: consideration of rejection in classification models. *IEEE Trans Biomed Eng* **55,** 108–118 (2008).
74. Voss, A. *et al.* Linear and nonlinear heart rate variability risk stratification in heart failure patients. *Computers in Cardiology, 2008* 557–560 (2008).
75. Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A. & Kurths, J. Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. *Phys. Rev. E* **66,** 026702 (2002).
76. Zbilut, J. P., Thomasson, N. & Webber, C. L. Recurrence quantification analysis as a tool for nonlinear exploration of nonstationary cardiac signals. *Medical Engineering & Physics* **24,** 53–60 (2002).
77. Caminal, P. *et al.* Symbolic Dynamic Analysis of Relations Between Cardiac and Breathing Cycles in Patients on Weaning Trials. *Ann Biomed Eng* **38,** 2542–2552 (2010).
78. Shiau, Y.-H. Detecting Well-Harmonized Homeostasis in Heart Rate Fluctuations. in *2008 International Conference on BioMedical Engineering and Informatics* 399–403 (2008). doi:10.1109/BMEI.2008.333
79. Javed, F. *et al.* Frequency spectrum analysis of finger photoplethysmographic waveform variability during haemodialysis. *Physiol Meas* **31,** 1203–1216 (2010).
80. Signal Processing Toolbox User's Guide - Version 6.14. (2010). at <http://www.mathworks.com/help/toolbox/signal/signal_product_page.html>
81. Liu, C., Liu, C., Li, L., Zhang, Q. & Li, B. Systolic and Diastolic Time Interval Variability Analysis and Their Relations with Heart Rate Variability. *Bioinformatics and Biomedical Engineering , 2009. ICBBE 2009. 3rd International Conference on* 1–4 (2009).

82. Valencia, J. F. *et al.* Complexity of the short-term heart-rate variability. *IEEE Eng Med Biol Mag* **28,** 72–78 (2009).
83. Kim, K. K., Kim, J. S., Lim, Y. G. & Park, K. S. The effect of missing RR-interval data on heart rate variability analysis in the frequency domain. *Physiol. Meas.* **30,** 1039–1050 (2009).
84. Roopaei, M., Boostani, R., Sarvestani, R. R., Taghavi, M. A. & Azimifar, Z. Chaotic based reconstructed phase space features for detecting ventricular fibrillation. *Biomedical Signal Processing and Control* **5,** 318–327 (2010).
85. Chua, K. C., Chandran, V., Acharya, U. R. & Lim, C. M. Application of higher order statistics/spectra in biomedical signals--A review. *Medical Engineering & Physics* **32,** 679–689 (2010).
86. Nikias, C. L. & Raghuveer, M. R. Bispectrum estimation: A digital signal processing framework. *Proceedings of the IEEE* **75,** 869–891 (2005).
87. Mateo, J. & Laguna, P. Improved heart rate variability signal analysis from the beat occurrence times according to the IPFM model. *IEEE Trans Biomed Eng* **47,** 985–996 (2000).
88. Martínez, J. P. *et al.* Detection Performance and Risk Stratification Using a Model-Based Shape Index Characterizing Heart Rate Turbulence. *Ann Biomed Eng* **38,** 3173–3184 (2010).
89. Laguna, P., Moody, G. B. & Mark, R. G. Power spectral density of unevenly sampled data by least-square analysis: performance and application to heart rate signals. *Biomedical Engineering, IEEE Transactions on* **45,** 698–715 (1998).
90. Muthuraman, M., Galka, A., Deuschl, G., Heute, U. & Raethjen, J. Dynamical correlation of non-stationary signals in time domain--A comparative study. *Biomedical Signal Processing and Control* **5,** 205–213 (2010).
91. Mitra, P. P. & Pesaran, B. Analysis of Dynamic Brain Imaging Data. *Biophysical Journal* **76,** 691–708 (1999).
92. Kantz, H. & Schreiber, T. *Nonlinear time series analysis*. (Cambridge Univ Pr, 2004).
93. Barbieri, R., Matten, E. C., Alabi, A. A. & Brown, E. N. A point-process model of human heartbeat intervals: new definitions of heart rate and heart rate variability. *Am J Physiol Heart Circ Physiol* **288,** H424–435 (2005).
94. Chen, Z., Brown, E. N. & Barbieri, R. Characterizing nonlinear heartbeat dynamics within a point process framework. *IEEE Trans Biomed Eng* **57,** 1335–1347 (2010).
95. Teich, M. C., Lowen, S. B., Jost, B. M. & others. *Heart Rate Variability: Measures and Models. Nonlinear Biomed Sig Proc Vol. II, Dynamic Analysis and Modeling*. (IEEE Press, New York, 2001).
96. Marwan, N., Carmen Romano, M., Thiel, M. & Kurths, J. Recurrence plots for the analysis of complex systems. *Physics Reports* **438,** 237–329 (2007).
97. Fernández, J. R., Hermida, R. C. & Mojón, A. Chronobiological analysis techniques. Application to blood pressure. *Philos Transact A Math Phys Eng Sci* **367,** 431–445 (2009).

98. Martinmaki, K., Rusko, H., Saalasti, S. & Kettunen, J. Ability of short-time Fourier transform method to detect transient changes in vagal effects on hearts: a pharmacological blocking study. *Am J Physiol Heart Circ Physiol* **290,** H2582–2589 (2006).
99. Sadiq, I. & Khan, S. A. Fuzzification of the Analysis of Heart Rate Variability Using ECG in Time, Frequency and Statistical Domains. in *Computer Engineering and Applications, International Conference on* **1,** 481–485 (IEEE Computer Society, 2010).
100. Voss, A. *et al.* Comparison of nonlinear methods symbolic dynamics, detrended fluctuation, and Poincaré plot analysis in risk stratification in patients with dilated cardiomyopathy. *Chaos* **17,** 015120 (2007).
101. Guzzetti, S. *et al.* Symbolic Dynamics of Heart Rate Variability: A Probe to Investigate Cardiac Autonomic Modulation. *Circulation* **112,** 465–470 (2005).
102. Maestri, R. *et al.* Nonlinear Indices of Heart Rate Variability in Chronic Heart Failure Patients: Redundancy and Comparative Clinical Value. *J Cardiovasc Electrophysiol* **18,** 425–433 (2007).
103. Wessel, N. *et al.* Short-term forecasting of life-threatening cardiac arrhythmias based on symbolic dynamics and finite-time growth rates. *Physical Review E* **61,** 733–739 (2000).
104. TORRENCE, C. & COMPO, G. A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society* **79,** 61–78 (1998).
105. Shafqat, K., Pal, S. K., Kumari, S. & Kyriacou, P. A. Time-frequency analysis of HRV data from locally anesthetized patients. *Conf Proc IEEE Eng Med Biol Soc* **2009,** 1824–1827 (2009).
106. Thurner, S. *et al.* Analysis, Synthesis, and Estimation of Fractal-Rate Stochastic Point Processes. *adap-org/9709006* (1997). at <http://arxiv.org/abs/adap-org/9709006>
107. Jovic, A. & Bogunovic, N. Random Forest-Based Classification of Heart Rate Variability Signals by Using Combinations of Linear and Nonlinear Features. in *XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010* 29–32 (2010).
108. Jović, A. & Bogunović, N. Feature Set Extension for Heart Rate Variability Analysis by Using Non-linear, Statistical and Geometric Measures. in *Proc. 31st Int. Conf. Inf. Tech. Interfaces, ITI* 35–40 (2009).
109. Liu, J., Zhang, C. & Zheng, C. EEG-based estimation of mental fatigue by using KPCA-HMM and complexity parameters. *Biomedical Signal Processing and Control* **5,** 124–130 (2010).
110. Ramdani, S., Seigle, B., Lagarde, J., Bouchara, F. & Bernard, P. L. On the use of sample entropy to analyze human postural sway data. *Medical engineering & physics* **31,** 1023–1031 (2009).
111. Chen, W., Zhuang, J., Yu, W. & Wang, Z. Measuring complexity using FuzzyEn, ApEn, and SampEn. *Medical Engineering & Physics* **31,** 61–68 (2009).

112. Baumert, M. *et al.* Forecasting of life threatening arrhythmias using the compression entropy of heart rate. *Methods of Information in Medicine* **43,** 202–206 (2004).
113. Liau, B.-Y., Yeh, S.-J., Chiu, C.-C. & Tsai, Y.-C. Dynamic cerebral autoregulation assessment using chaotic analysis in diabetic autonomic neuropathy. *Med Bio Eng Comput* **46,** 1–9 (2007).
114. Castiglioni, P., Parati, G., Civijian, A., Quintin, L. & Di Rienzo, M. Local scale exponents of blood pressure and heart rate variability by detrended fluctuation analysis: effects of posture, exercise, and aging. *IEEE Trans Biomed Eng* **56,** 675–684 (2009).
115. Schmitt, D. T., Stein, P. K. & Ivanov, P. C. Stratification pattern of static and scale-invariant dynamic measures of heartbeat fluctuations across sleep stages in young and elderly. *IEEE Trans Biomed Eng* **56,** 1564–1573 (2009).
116. Delignieres, D. *et al.* Fractal analyses for 'short' time series: A re-assessment of classical methods. *Journal of Mathematical Psychology* **50,** 525–544 (2006).
117. Allegrini, P., Grigolini, P., Hamilton, P., Palatella, L. & Raffaelli, G. Memory beyond memory in heart beating, a sign of a healthy physiological condition. *Phys. Rev. E* **65,** 041926 (2002).
118. Allegrini, P. *et al.* Long- and short-time analysis of heartbeat sequences: Correlation with mortality risk in congestive heart failure patients. *Phys. Rev. E* **67,** 062901 (2003).
119. Grigolini, P., Palatella, L. & Raffaelli, G. Asymmetric anomalous diffusion: An efficient way to detect memory in time series. *Fractals* **9,** 439–449 (2001).
120. Fichera, A., Losenno, C. & Pagano, A. Clustering of chaotic dynamics of a lean gas-turbine combustor. *Applied Energy* **69,** 101–117 (2001).
121. Michieli, I., Medved, B. & Ristov, S. Data series embedding and scale invariant statistics. *Hum Mov Sci* **29,** 449–463 (2010).
122. Agarwal, R., Gotman, J., Flanagan, D. & Rosenblatt, B. Automatic EEG analysis during long-term monitoring in the ICU. *Electroencephalography and Clinical Neurophysiology* **107,** 44–58 (1998).
123. Ruffo, M., Cesarelli, M., Romano, M., Bifulco, P. & Fratini, A. An algorithm for FHR estimation from foetal phonocardiographic signals. *Biomedical Signal Processing and Control* **5,** 131–141 (2010).
124. Rangayyan, R. M. & Wu, Y. Analysis of vibroarthrographic signals with features related to signal variability and radial-basis functions. *Ann Biomed Eng* **37,** 156–163 (2009).
125. Patangay, A., Zhang, Y. & Lewicke, A. Measures of cardiac contractility variability during ischemia. *Conf Proc IEEE Eng Med Biol Soc* **2009,** 4198–4201 (2009).
126. Amann, A., Tratnig, R. & Unterkofler, K. Detecting ventricular fibrillation by time-delay methods. *IEEE Trans Biomed Eng* **54,** 174–177 (2007).

127. Bauer, A. *et al.* Heart Rate Turbulence: Standards of Measurement, Physiological Interpretation, and Clinical Use: International Society for Holter and Noninvasive Electrophysiology Consensus. *J Am Coll Cardiol* **52,** 1353–1365 (2008).
128. Affinito, M., Carrozzi, M., Accardo, A. & Bouquet, F. Use of the fractal dimension for the analysis of electroencephalographic time series. *Biological Cybernetics* **77,** 339–350 (1997).
129. Frank, B., Pompe, B., Schneider, U. & Hoyer, D. Permutation entropy improves fetal behavioural state classification based on heart rate analysis from biomagnetic recordings in near term fetuses. *Med Bio Eng Comput* **44,** 179–187 (2006).
130. Hu, J., Gao, J. & Tung, W. Characterizing heart rate variability by scale-dependent Lyapunov exponent. *Chaos* **19,** 028506 (2009).
131. Myers, S. A., Stergiou, N., Pipinos, I. I. & Johanning, J. M. Gait Variability Patterns are Altered in Healthy Young Individuals During the Acute Reperfusion Phase of Ischemia-Reperfusion. *Journal of Surgical Research* **164,** 6–12 (2010).
132. De la Vega, D. *et al.* Generalization of Lee Method for the analysis of the signal variability. *IEEE transactions on vehicular technology* **58,** 506–516 (2009).
133. Ivanov, P. C. *et al.* Multifractality in human heartbeat dynamics. *Nature* **399,** 461–465 (1999).
134. Abry, P. *et al.* Multifractal analysis of ECG for intrapartum diagnosis of fetal asphyxia. in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on* 566–569 (2010).
135. Costa, M., Goldberger, A. L. & Peng, C.-K. Multiscale entropy analysis of biological signals. *Phys Rev E Stat Nonlin Soft Matter Phys* **71,** 021906 (2005).
136. Costa, M. D., Peng, C.-K. & Goldberger, A. L. Multiscale analysis of heart rate dynamics: entropy and time irreversibility measures. *Cardiovasc Eng* **8,** 88–93 (2008).
137. Porta, A. *et al.* Complexity and nonlinearity in short-term heart period variability: comparison of methods based on local nonlinear prediction. *IEEE Trans Biomed Eng* **54,** 94–106 (2007).
138. Kampouraki, A., Manis, G. & Nikou, C. Heartbeat time series classification with support vector machines. *IEEE Trans Inf Technol Biomed* **13,** 512–518 (2009).
139. Porta, A. *et al.* Prediction of short cardiovascular variability signals based on conditional distribution. *Biomedical Engineering, IEEE Transactions on* **47,** 1555–1564 (2002).
140. Javorka, M. *et al.* Recurrences in heart rate dynamics are changed in patients with diabetes mellitus. *Clinical Physiology and Functional Imaging* **28,** 326–331 (2008).
141. Schmit, J. M. *et al.* Deterministic center of pressure patterns characterize postural instability in Parkinson's disease. *Exp Brain Res* **168,** 357–367 (2005).
142. Webber Jr, C. L. & Zbilut, J. P. Dynamical assessment of physiological systems and states using recurrence plot strategies. *Journal of Applied Physiology* **76,** 965 (1994).
143. Huang, H.-H. *et al.* Using a short-term parameter of heart rate variability to distinguish awake from isoflurane anesthetic states. *Med Biol Eng Comput* **46,** 977–984 (2008).

144. Yang, A. C.-C., Hseu, S.-S., Yien, H.-W., Goldberger, A. L. & Peng, C.-K. Linguistic Analysis of the Human Heartbeat Using Frequency and Rank Order Statistics. *Phys. Rev. Lett.* **90,** 108103 (2003).
145. Jovic, A. & Bogunovic, N. Feature extraction for ECG time-series mining based on chaos theory. in *Information Technology Interfaces, 2007. ITI 2007. 29th International Conference on* 63–68 (2007).
146. Thamrin, C. *et al.* Predicting future risk of asthma exacerbations using individual conditional probabilities. *J. Allergy Clin. Immunol.* **127,** 1494–1502.e3 (2011).
147. Frey, U. *et al.* Risk of severe asthma episodes predicted from fluctuation analysis of airway function. *Nature* **438,** 667–670 (2005).
148. Mendonca, G. V., Fernhall, B., Heffernan, K. S. & Pereira, F. D. Spectral methods of heart rate variability analysis during dynamic exercise. *Clin Auton Res* **19,** 237–245 (2009).
149. Arnold, R., Lundy, D., Glaspey, L., Green, G. & Seely, A. Heart rate variability in the early resuscitation of septic shock. *Crit Care* **13,** P50–P50 (2009).
150. Chen, W., Wang, Z., Xie, H. & Yu, W. Characterization of surface EMG signal based on fuzzy entropy. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on* **15,** 266–272 (2007).
151. Whitacre, J. M. Degeneracy: a link between evolvability, robustness and complexity in biological systems. *Theor Biol Med Model* **7,** 6 (2010).
152. Glass, L. Synchronization and rhythmic processes in physiology. *Nature* **410,** 277–284 (2001).
153. Schafer, C., Rosenblum, M. G., Kurths, J. & Abel, H.-H. Heartbeat synchronized with ventilation. *Nature* **392,** 239–240 (1998).
154. Huikuri, H. V., Perkiömäki, J. S., Maestri, R. & Pinna, G. D. Clinical impact of evaluation of cardiovascular control by novel methods of heart rate dynamics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **367,** 1223 –1238 (2009).
155. Voss, A., Schulz, S., Schroeder, R., Baumert, M. & Caminal, P. Methods derived from nonlinear dynamics for analysing heart rate variability. *Philos Transact A Math Phys Eng Sci* **367,** 277–296 (2009).
156. Aletti, F., Bassani, T., Lucini, D., Pagani, M. & Baselli, G. Multivariate decomposition of arterial blood pressure variability for the assessment of arterial control of circulation. *IEEE Trans Biomed Eng* **56,** 1781–1790 (2009).
157. Migeotte, P. F. & Verbandt, Y. A novel algorithm for the heart rate variability analysis of short-term recordings: polar representation of respiratory sinus arrhythmia. *Comput. Biomed. Res* **32,** 56–66 (1999).
158. Colombo, J. *et al.* Noninvasive monitoring of the autonomic nervous system and hemodynamics of patients with blunt and penetrating trauma. *J Trauma* **65,** 1364–1373 (2008).
159. Nozawa, M., Yana, K., Kaeriyama, K., Mizuta, H. & Ono, T. Spontaneous variability analysis for characterizing cardiovascular responses to water ingestion. *Conf Proc IEEE Eng Med Biol Soc* **2009,** 1816–1819 (2009).

160. Shaffer, J. P. Multiple hypothesis testing. *Annual review of psychology* **46,** 561–584 (1995).
161. Farcomeni, A. A review of modern multiple hypothesis testing, with particular attention to the false discovery proportion. *Stat Methods Med Res* **17,** 347–388 (2008).
162. Hastie, T., Tibshirani, R. & Friedman, J. J. H. *The elements of statistical learning*. **1,** (Springer New York, 2001).
163. Bishop, C. M. & Nasrabadi, N. M. *Pattern recognition and machine learning*. **1,** (springer New York, 2006).
164. Guyon, I. & Elisseeff, A. An introduction to variable and feature selection. *The Journal of Machine Learning Research* **3,** 1157–1182 (2003).
165. Hyvärinen, A. & Oja, E. Independent component analysis: algorithms and applications. *Neural networks* **13,** 411–430 (2000).
166. Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. & Doyne Farmer, J. Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena* **58,** 77–94 (1992).
167. Theodoridis, S. & Koutroumbas, K. *Pattern recognition*. (Academic Press, 2009).
168. Kotani, K., Struzik, Z. R., Takamasu, K., Stanley, H. E. & Yamamoto, Y. Model for complex heart rate dynamics in health and diseases. *Phys. Rev. E* **72,** 041904 (2005).
169. Seidel, H. & Herzel, H. Bifurcations in a nonlinear model of the baroreceptor-cardiac reflex. *Physica D: Nonlinear Phenomena* **115,** 145–160 (1998).
170. Kotani, K., Takamasu, K., Ashkenazy, Y., Stanley, H. E. & Yamamoto, Y. Model for cardiorespiratory synchronization in humans. *PHYSICAL REVIEW-SERIES E-* **65,** 51923–51923 (2002).
171. Dudkowska, A. & Makowiec, D. Seidel–Herzel model of human baroreflex in cardiorespiratory system with stochastic delays. *Journal of mathematical biology* **57,** 111–137 (2008).
172. Duggento, A., Toschi, N. & Guerrisi, M. Modeling of human baroreflex: considerations on the Seidel-Herzel model. *Fluctuation and Noise Letters* **11,** (2012).
173. Bravi, A. CIMVA Core Description manual - http://ohridal.org/cimva/CIMVA-Core-Description.pdf. (2013). at <http://ohridal.org/cimva/CIMVA-Core-Description.pdf>
174. Vinh, N. X., Epps, J. & Bailey, J. Information Theoretic Measures for Clusterings Comparison: Variants, Properties, Normalization and Correction for Chance. *The Journal of Machine Learning Research* **9999,** 2837–2854 (2010).
175. Botev, Z. I., Grotowski, J. F. & Kroese, D. P. Kernel density estimation via diffusion. *The Annals of Statistics* **38,** 2916–2957 (2010).
176. Schaffernicht, E., Kaltenhaeuser, R., Verma, S. S. & Gross, H.-M. in *Artificial Neural Networks – ICANN 2010* (Diamantaras, K., Duch, W. & Iliadis, L. S.) **6352,** 362–367 (Springer Berlin Heidelberg, 2010).
177. Tan, C. O. Heart rate variability: are there complex patterns? *Frontiers in Physiology* **4,** 165 (2013).

178. Tan, C. O., Cohen, M. A., Eckberg, D. L. & Taylor, J. A. Fractal properties of human heart period variability: physiological and methodological implications. *The Journal of physiology* **587,** 3929–3941 (2009).
179. Seely, A. E., Bravi, A. & Longtin, A. System and Method for Generating Composite Measures of Variability. (2013). at <http://patentscope.wipo.int/search/en/WO2013003953>
180. Bravi, A., Green, G., Longtin, A. & Seely, A. J. Monitoring and Identification of Sepsis Development through a Composite Measure of Heart Rate Variability. *PLOS ONE* **7,** e45666 (2012).
181. Lewis, M. J. & Short, A. L. Exercise and cardiac regulation: what can electrocardiographic time series tell us? *Scandinavian Journal of Medicine & Science in Sports* **20,** 794–804 (2010).
182. Perini, R. & Veicsteinas, A. Heart rate variability and autonomic activity at rest and during exercise in various physiological conditions. *European Journal of Applied Physiology* **90,** 317–325 (2003).
183. Routledge, F. S., Campbell, T. S., McFetridge-Durdle, J. A. & Bacon, S. L. Improvements in heart rate variability with exercise therapy. *Can J Cardiol* **26,** 303–312 (2010).
184. Varela, M., Ruiz-Esteban, R. & Mestre de Juan, M. J. Chaos, fractals, and our concept of disease. *Perspect. Biol. Med.* **53,** 584–595 (2010).
185. Goldberger, A. L. Non-linear dynamics for clinicians: chaos theory, fractals, and complexity at the bedside. *The Lancet* **347,** 1312–1314 (1996).
186. Barrera-Ramirez, J., Bravi, A., Geoffrey, G., Seely, A. & Kenny, G. Comparison of Heart and Respiratory Rate Variability Measures Using an Intermittent Incremental Submaximal Exercise Model. *Applied Physiology, Nutrition, and Metabolism* **ID: apnm-2012-0370.R3,** (2013).
187. Clifford, G. D., McSharry, P. E. & Tarassenko, L. Characterizing artefact in the normal human 24-hour RR time series to aid identification and artificial replication of circadian variations in human beat to beat heart rate using a simple threshold. in *Computers in Cardiology, 2002* 129–132 (2002). at <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1166724>
188. Shephard, R. J. Sepsis and mechanisms of inflammatory response: is exercise a good model? *Br J Sports Med* **35,** 223–230 (2001).
189. Bente Klarlund, P. in *Natural Immunity* **Volume 5,** 263–288 (Elsevier, 2005).
190. Vincent, J.-L. Understanding cardiac output. *Critical Care* **12,** 174 (2008).
191. Buchan, C., Bravi, A. & Seely, A. Variability Analysis and the Diagnosis, Management, and Treatment of Sepsis. *Current Infectious Disease Reports* **14,** 512–521 (2012).
192. Chen, W. L. & Kuo, C. D. Characteristics of heart rate variability can predict impending septic shock in emergency department patients with sepsis. *Academic emergency medicine* **14,** 392–397 (2007).

193. Musen, M. A., Shahar, Y. & Shortliffe, E. H. in *Biomedical informatics* 698–736 (Springer, 2006). at <http://link.springer.com/content/pdf/10.1007/0-387-36278-9_20.pdf>
194. Phansalkar, S. *et al.* A review of human factors principles for the design and implementation of medication safety alerts in clinical information systems. *J Am Med Inform Assoc* **17,** 493–501 (2010).
195. Yen, P.-Y. & Bakken, S. Review of health information technology usability study methodologies. *J Am Med Inform Assoc* **19,** 413–422 (2012).
196. Brindley, P. G., Zhu, N. & Sligl, W. Best evidence in critical care medicine Early antibiotics and survival from septic shock: it's about time. *Can J Anesth/J Can Anesth* **53,** 1157–1160 (2006).
197. Angus, D. C. *et al.* Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. *Crit. Care Med* **29,** 1303–1310 (2001).
198. Moorman, J. R., Lake, D. E. & Griffin, M. P. Heart rate characteristics monitoring for neonatal sepsis. *Biomedical Engineering, IEEE Transactions on* **53,** 126–132 (2006).
199. Griffin, M. P. *et al.* Heart Rate Characteristics: Novel Physiomarkers to Predict Neonatal Infection and Death. *Pediatrics* **116,** 1070 –1074 (2005).
200. Zimmerman, J. L. *et al.* Surviving Sepsis Campaign guidelines for management of severe sepsis and septic shock. *Intensive Care Medicine* **30,** 536–555 (2004).
201. Nguyen, H. *et al.* Severe Sepsis and Septic Shock: Review of the Literature and Emergency Department Management Guidelines. *Annals of Emergency Medicine* **48,** 54.e1–54.e1 (2006).
202. Ahmad, S. *et al.* Measurement of Heart Rate Variability Using an Oscillometric Blood Pressure Monitor. *IEEE Transactions on Instrumentation and Measurement* **59,** 2575–2590 (2010).
203. Ahmad, S., Tejuja, A., Newman, K., Zarychanski, R. & Seely, A. Clinical review: A review and analysis of heart rate variability and the diagnosis and prognosis of infection. *Critical Care* **13,** 232 (2009).
204. Saria, S., Rajani, A. K., Gould, J., Koller, D. & Penn, A. A. Integration of Early Physiological Responses Predicts Later Illness Severity in Preterm Infants. *Sci Transl Med* **2,** 48ra65–48ra65 (2010).
205. Günther, A. *et al.* Heart rate variability – a potential early marker of sub-acute post-stroke infections. *Acta Neurologica Scandinavica* doi:10.1111/j.1600-0404.2011.01626.x
206. De Jonghe, B. *et al.* Respiratory weakness is associated with limb weakness and delayed weaning in critical illness*. *Critical Care Medicine* **35,** 2007–2015 (2007).
207. Rello, J. *et al.* Epidemiology and outcomes of ventilator-associated pneumonia in a large US database. *Chest* **122,** 2115–2121 (2002).
208. Rajakaruna, C., Rogers, C. A., Angelini, G. D. & Ascione, R. Risk factors for and economic implications of prolonged ventilation after cardiac surgery. *J. Thorac. Cardiovasc. Surg.* **130,** 1270–1277 (2005).

209. Epstein, S. K., Ciubotaru, R. L. & Wong, J. B. Effect of failed extubation on the outcome of mechanical ventilation. *Chest* **112,** 186–192 (1997).
210. Esteban, A. *et al.* A comparison of four methods of weaning patients from mechanical ventilation. *New England Journal of Medicine* **332,** 345–350 (1995).
211. Epstein, S. K. Predicting Extubation Failure Is It in (on) the Cards? *CHEST Journal* **120,** 1061–1063 (2001).
212. Epstein, S. K., Zilberberg, M. D., Jacoby, C., Ciubotaru, R. L. & Kaplan, L. M. Response to symptom-limited exercise in patients with the hepatopulmonary syndrome. *CHEST Journal* **114,** 736–741 (1998).
213. Needham, D. M. & Pronovost, P. J. The importance of understanding the costs of critical care and mechanical ventilation*. *Critical care medicine* **33,** 1434–1435 (2005).
214. Dasta, J. F., McLaughlin, T. P., Mody, S. H. & Piech, C. T. Daily cost of an intensive care unit day: The contribution of mechanical ventilation*. *Critical care medicine* **33,** 1266–1271 (2005).
215. Zeggwagh, A. A., Abouqal, R., Madani, N., Zekraoui, A. & Kerkeb, O. Weaning from mechanical ventilation: a model for extubation. *Intensive care medicine* **25,** 1077–1083 (1999).
216. Yang, K. L. & Tobin, M. J. A prospective study of indexes predicting the outcome of trials of weaning from mechanical ventilation. *New England Journal of Medicine* **324,** 1445–1450 (1991).
217. Meade, M. *et al.* Predicting success in weaning from mechanical ventilation. *Chest* **120,** 400S–24S (2001).
218. Frutos-Vivar, F. *et al.* Risk factors for extubation failure in patients following a successful spontaneous breathing trial. *CHEST Journal* **130,** 1664–1671 (2006).
219. MacIntyre, N. Discontinuing mechanical ventilatory support. *Chest* **132,** 1049–1056 (2007).
220. Esteban, A. *et al.* Effect of spontaneous breathing trial duration on outcome of attempts to discontinue mechanical ventilation. *American journal of respiratory and critical care medicine* **159,** 512–518 (1999).
221. Thille, A. W., Richard, J.-C. M. & Brochard, L. The Decision to Extubate in the Intensive Care Unit. *American journal of respiratory and critical care medicine* (2013). at <http://www.atsjournals.org/doi/abs/10.1164/rccm.201208-1523CI>
222. Jacono, F. J. & Dick, T. E. Variability, measuring the spice of life. *J. Appl. Physiol.* **111,** 351–352 (2011).
223. Bien, M.-Y. *et al.* Breathing pattern variability: a weaning predictor in postoperative patients recovering from systemic inflammatory response syndrome. *Intensive Care Med* **30,** 241–247 (2004).
224. Hamilton, P. S. & Tompkins, W. J. Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. *Biomedical Engineering, IEEE Transactions on* 1157–1165 (1986).
225. Clifford, G. D., McSharry, P. E. & Tarassenko, L. Characterizing artefact in the normal human 24-hour RR time series to aid identification and artificial replication of

circadian variations in human beat to beat heart rate using a simple threshold. in *Computers in Cardiology, 2002* 129–132 (2002). at <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1166724>

226. Li, Q., Mark, R. G. & Clifford, G. D. Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. *Physiological measurement* **29,** 15 (2007).
227. Bhavani-Shankar, K. & Philip, J. H. Defining segments and phases of a time capnogram. *Anesthesia & Analgesia* **91,** 973–977 (2000).
228. Pounds, S. & Cheng, C. Robust estimation of the false discovery rate. *Bioinformatics* **22,** 1979 –1987 (2006).
229. Polikar, R. Ensemble based systems in decision making. *IEEE Circuits and Systems Magazine* **6,** 21 –45 (2006).
230. Boles, J.-M. *et al.* Weaning from mechanical ventilation. *European Respiratory Journal* **29,** 1033–1056 (2007).
231. Mehta, S., Nelson, D. L., Klinger, J. R., Buczko, G. B. & Levy, M. M. Prediction of post-extubation work of breathing. *Crit. Care Med.* **28,** 1341–1346 (2000).
232. El-Khatib, M., Jamaleddine, G., Soubra, R. & Muallem, M. Pattern of spontaneous breathing: potential marker for weaning outcome. Spontaneous breathing pattern and weaning from mechanical ventilation. *Intensive Care Med* **27,** 52–58 (2001).
233. Papaioannou, V. E., Chouvarda, I., Maglaveras, N., Dragoumanis, C. & Pneumatikos, I. Changes of heart and respiratory rate dynamics during weaning from mechanical ventilation: A study of physiologic complexity in surgical critically ill patients. *Journal of Critical Care* **26,** 262–272 (2011).
234. Arcentales, A., Giraldo, B. F., Caminal, P., Benito, S. & Voss, A. Recurrence quantification analysis of heart rate variability and respiratory flow series in patients on weaning trials. in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society,EMBC* 2724–2727 (2011). doi:10.1109/IEMBS.2011.6090747
235. Clifford, G. D. & others. ECG statistics, noise, artifacts, and missing data. *Advanced Methods and Tools for ECG Analysis* (2006). at <http://www.robots.ox.ac.uk/~gari/ecgbook/ch3.pdf>
236. Clifford, G. D., Behar, J., Li, Q. & Rezek, I. Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms. *Physiological measurement* **33,** 1419 (2012).
237. Bien, M.-Y. *et al.* Comparisons of predictive performance of breathing pattern variability measured during T-piece, automatic tube compensation, and pressure support ventilation for weaning intensive care unit patients from mechanical ventilation*. *Critical care medicine* **39,** 2253–2262 (2011).
238. Foteinou, P. T., Calvano, S. E., Lowry, S. F. & Androulakis, I. P. A physiological model for autonomic heart rate regulation in human endotoxemia. *Shock (Augusta, Ga.)* **35,** 229 (2011).

239. Scheff, J. D., Mavroudis, P. D., Calvano, S. E., Lowry, S. F. & Androulakis, I. P. Modeling autonomic regulation of cardiac function and heart rate variability in human endotoxemia. *Physiological genomics* **43,** 951–964 (2011).
240. Namas, R. *et al.* Sepsis: Something old, something new, and a systems view. *Journal of Critical Care* doi:10.1016/j.jcrc.2011.05.025
241. Breiman, L. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science* **16,** 199–231 (2001).