ASSESSMENT OF VENTILATORY THRESHOLDS FROM HEART RATE VARIABILITY IN FIVE INCREMENTAL TREADMILL TESTS IN CROSS COUNTRY SKIERS

Ibai Mendia Iztueta

Master's Thesis in Exercise Physiology
Spring 2014
Department of Biology of Physical Activity
University of Jyväskylä

ABSTRACT

Mendia Iztueta, Ibai 2014. Assessment of Ventilatory Thresholds from Heart Rate Variability in Five Incremental Treadmill Tests in Cross Country Skiers. Department of Biology of Physical Activity, University of Jyväskylä. Master's Thesis in Exercise Physiology. 103 pp.

Incremental treadmill tests are widely used in the field of exercise physiology for the assessment of Ventilatory Thresholds for clinical and sport oriented issues. The assessment of Ventilatory Thresholds (VTs) from Heart Rate Variability (HRV) is a relatively new approach with increasing popularity because it is a non-invasive and economical method. Nevertheless, this has not been used in Cross Country (XC) Skiing, an endurance sport where the knowledge of VTs holds special importance. The purpose of the study was to assess VTs using the data derived from HRV analysis in the five main XC skiing-related techniques, Double Poling (DP), Diagonal Striding (DS), Nordic Walking (NW), V1 Skating (V1), and V2 Skating (V2). V1 consists of an slightly off-set double poling action on every other skating glide, which is used for hill climbing, whereas the V2 technique uses a double poling action on every skating leg and is used on flat and on moderate uphills. Ten competitive national level skiers completed the five incremental treadmill tests until exhaustion with a minimum of one to two recovery days in between each test. Ventilatory gases, HRV and Poling Frequency (PF) were measured. Two methods ($HRVT_{1\text{-}STD}$ and $HRVT_{1\text{-}MSD}$) derived from time-domain analyses of HRV were used for the assessment of the first VT (VT_1), whereas the second VT (VT_2) was assessed with two frequency analyses methods ($HRVT_{2\text{-}HFP}$ and $HRVT_{2\text{-}HFP\text{-}RSA}$). The results showed that the only cases where the proposed HRVTs (i.e. Heart Rate Variability Thresholds) were good assessors of their respective VTs were the $HRVT_{1\text{-}STD}$ in the DS test, the $HRVT_{1\text{-}MSD}$ in the DS and V2 tests, and the $HRVT_{2\text{-}HFP\text{-}RSA}$ in the NW test. The lack of a wider success of the assessment of VTs from HRV was reasoned to be, among other things, due to the high entrainment between the Breathing Frequency (BF) and PF, where the rhythm of PF alters the usual ventilatory response in which the assessment of both VTs and HRVTs is based.

Keywords: Exercise Physiology, Cross Country Ski, Ventilatory Thresholds, Heart Rate Variability, HRV, Poling Frequency, Breathing Frequency, Assessment.

ACKNOWLEDGEMENTS

The present study was made in cooperation between the Department of Biology of Physical Activity (DBPA) at the University of Jyväskylä and the Research Institute for Olympic Sports (KIHU, initials used from the Finnish language) as a part of a broader study which produced another Master’s thesis. The measurements were carried out in the laboratory of KIHU.

I would like to acknowledge both institutions and in particular my supervisors Doctor Esa Hynynen (KIHU) and Professor Heikki Kyröläinen (DBPA), also Kristen Monahan (DBPA), co-worker who produced the other Master's thesis. I also wish to express my biggest gratitude to the company Firstbeat Technologies Oy for offering their software for the heart rate analysis, and also to the Department of Education, Linguistic Politics and Culture of the Basque Government for their financial support.

I would also like to express my most sincere appreciation to all the people who, in one way or another, were involved in the research project, and who thus made the current thesis possible. Among these, I want to give special thanks to Sirpa Vänttinen, employee of KIHU, for her support during the measurements, and also to the members of the DBPA; the professors Heikki Kainulainen, Antti Mero, and Teemu Pullinen, and my classmates. I do not want to forget to acknowledge the help that I received concerning technical aspects, such as statistical analysis from other staff members and students from the University of Jyväskylä, Svenja Wachsmuth and professor Timo Törmäkangas. Nevertheless, the support that I received from outside Jyväskylä regarding the spectral analysis has also been indispensable, and I would particularly like to express my gratitude to professor Rocco Di Michele and Laurent Mourot.

Finally, my appreciation also goes out to the more than talented subjects, without their effort the study would have not been possible, and, also to my family and girlfriend, Marianne, whose continuous love and support towards my studies have been indispensable.

LIST OF ABBREVIATIONS

- ANS - Autonomic Nervous System
- aRSA - Amplitude of RSA
- BF - Breathing Frequency
- DP - Double Poling
- DS - Diagonal Striding
- fHF - Dominant frequency in HF range
- HF - High Frequency
- HFP - HF Power
- HFP_{LOC} - LOC component of HFP
- HFP_{RSA} - RSA component of HFP
- HRV - Heart Rate Variability
- $HRVT_1$ - First HRV Threshold
- $HRVT_{1-MSD}$ - $HRVT_1$ determined from MSD trend
- $HRVT_{1-STD}$ - $HRVT_1$ determined from STD trend
- $HRVT_2$ - Second HRV Threshold
- $HRVT_{2-HFP}$ - $HRVT_2$ determined from HFP trend
- $HRVT_{2-HFP-RSA}$ - $HRVT_2$ determined from HFP_{RSA} trend
- LF - Low Frequency
- LFP - LF Power
- LOC - Cardiolocomotor Coupling
- LRC - Locomotor-Respiratory Coupling
- LT_1 - First Lactate Threshold
- LT_2 - Second Lactate Threshold
- MSD - Mean Successive Difference
- NW - Nordic Walking
- PF - Poling Frequency
- pfHF - Frequency corresponding to PF in HF range
- RSA - Respiratory Sinus Arrhythmia
- STD - Standard Deviation
- tV - Tidal Volume
- $\mathring{V}E$ - Minute Ventilation
- $\mathring{V}O_2$ - Oxygen Uptake
- $\mathring{V}CO_2$ - Carbon Dioxide Output
- VT_1 - First Ventilatory Threshold
- VT_2 - Second Ventilatory Threshold
- V1 - V1 skating
- V2 - V2 skating
- XC - Cross Country

CONTENTS

1 INTRODUCTION

Endurance exercise is a worldwide extended exercise mode that is expressed in multiple contexts with recreational, competitive, clinical or displacement purposes. The consciousness of endurance capacity is essential for the prescription and control of exercise as well as for the prediction of exercise performance. This requires cautious cardiopulmonary exercise testing that has been implemented in the area of Sport Science where competitive excellence is targeted by means such as optimization of endurance performance, but also in medicine, where endurance exercise has been implemented as a therapeutic measure in several cardiac and pulmonary disease entities, (e.g. obstructive airway disease, coronary artery disease, and chronic heart failure; Meyer et al. 2005).

Maximal oxygen uptake and maximal power output attained during incremental ergometric testing are the most frequently applied indicators of endurance capacity. However, within the last two decades, blood lactate curves and gas exchange measurements from incremental exercise tests have also been used for the assessment of endurance capacity, where the existence of two submaximal thresholds have been revealed (Meyer et al. 2005). These two submaximal thresholds present the advantage of predicting endurance capacity without the need for using strenuous nature of maximal exercise tests, often contraindicated for clinical and sporting reasons. In this study, the terms first and second Lactate Thresholds (LT_1 & LT_2) will be used to refer to the thresholds that have been assessed from the blood lactate concentration, whereas the terms first and second Ventilatory Thresholds (VT_1 & VT_2) will refer to the thresholds assessed from the ventilatory gas exchange data. The abbreviations LTs and VTs will refer to both first and second Lactate or Ventilatory Thresholds, respectively.

LT_1, corresponds to the first increase in blood lactate concentrations above resting values originating from muscle glycogenolysis, and is concomitant to the VT_1. Whereas LT_2 corresponds to VT_2, and approximates the Maximal Lactate Steady State, which represents the exercise intensity above which a continuous increase in blood lactate is unavoidable (Meyer et al. 2005). However, despite these physiological thresholds having a broad framework of applicability, their assessment either from blood lactate analysis or from ventilatory gas analysis, also present some limitations. The identification of the VTs depends on the use of sophisticated, expensive gas analyzers that are frequently laboratory-dependent. Although the analysis of lactate concentration is sometimes inexpensive, the successive use of the invasive nature of the blood collections makes this method impractical in many situations, especially in the clinical setting (Dourado et al. 2010).

At the beginning of the 21st century, attempts to overcome these limitations begun towards the development of a new methodology for the assessment of the physiological thresholds based on Heart Rate Variability (HRV). It was soon concluded that the use of HRV might be a simple and inexpensive alternative for the determination of VTs, with the additional benefits of being non-invasive and non-laboratory dependent. (e.g. Anosov et al. 2000; Cottin et al. 2004; Blain et al. 2005b.)

In this paper, the inflection points obtained from HRV signal curve will be referred to as the HRV Threshold 1 and 2 ($HRVT_1$ and $HRVT_2$), which the literature has shown to closely correlate with the VTs and LTs (e,g. Blain et al. 2005b; Cottin et al. 2006; Cottin et. al. 2007). Di Michele et al. (2012) clarifies that as the HRVTs (both $HRVT_1$ and $HRVT_2$) are obtained by estimating ventilatory parameters from HRV, it may be more consistent to compare them with the VTs, as done in previous studies (Blain et al. 2005b; Buchheit et al. 2007; Cottin et al. 2006 & 2007), rather than to the lactate-based gold standards. The purpose of the present study was to assess the VTs from HRV analysis using five incremental treadmill tests to exhaustion with competitive cross country (XC) skiers.

The selected techniques were based on the most common exercise modes for training in XC skiing: Diagonal Striding (DS), Nordic Walking (NW), Double Poling (DP), V1 skating (V1), and V2 skating (V2). DS is a running action with parallel skis, kicking and gliding with each stride, where the poles are planted alternately on the opposite side to the kick, which is mostly used on uphills and accelerating. NW is a total body version of walking where a bilateral asynchronous poling as well as a lengthened stride length are used to mimic the DS technique. DP is a discipline where the propulsion force is mostly provided by the simultaneous action of both poles; it is mostly used on flat, gentle climbing, and gentle downhills. V1 consists of an slightly off-set double poling action on every other skating glide, this technique is used for hill climbing. The V2 technique uses a double poling action on every skating leg; it is used on flat for accelerating or on moderate uphills.

2 LITERATURE REVIEW

2.1 Modulation of Heart Rate Variability

HRV is defined as the continuous beat-to-beat oscillation of R-R intervals around its mean value on the Electrocardiogram, measured over a period ranging from a few minutes to 24 hours (Freeman et al. 2006; Hynynen et al. 2006). An informative Electrocardiogram output can be observed in Figure 1, where the concept of HRV is clearly illustrated.

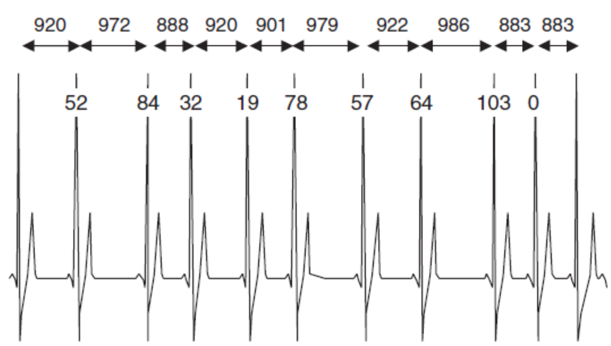

FIGURE 1. Example of an Electrocardiogram output over 11 beats where R-R interval times and difference between adjacent R-R intervals are displayed (Achten & Jeukendrup 2003).

This physiological secondary variation of Heart Rate (HR), which is characterised by a circadian variation, is related to changes in respiration (namely Respiratory Sinus Arrhythmia or RSA), but also to physical, environmental, mental and multiple other factors. Both the basic HR and its modulation are primarily controlled by higher brain centres and cardiovascular control areas in the brain stem through variations on the activity of the Autonomic Nervous System (ANS). (Aubert et al. 2003; Hainsworth 1998.)

The term ANS is attributed to J. N. Langley in the early 20th century to describe the nerves that are predominantly concerned with the regulation of bodily functions. These nerves generally function without consciousness, although this distinction from the Somatic Nervous System is not absolute. Autonomic nerves are divided into two divisions. The Sympathetic Nervous System with thoraco-lumbar outflow, and the preganglionic outflow of the Parasympathetic Nervous System, that arises from the brain stem that is known as the craniosacral outflow. Both divisions contain afferent and efferent nerves, and in a simplified way, it can be said that they mainly present complementary functions: the Sympathetic Nervous System excites the heart, constricting blood vessels, decreasing gastrointestinal motility and constricting sphincters, and the Parasympathetic Nervous System induces the opposite responses (Hainsworth 1998; Aubert et al. 2003). Consequently, the increased parasympathetic activity slows the HR, and the increased sympathetic activity increases the HR as it is shown in the Figure 2.

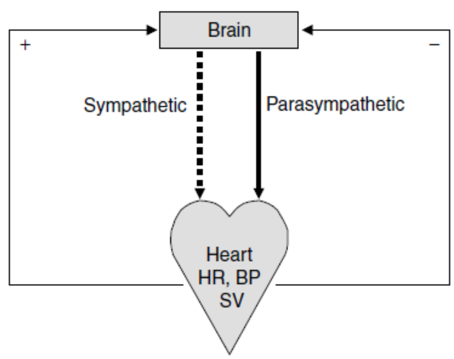

FIGURE 2. A simplified model illustrating the influence of the sympathetic and parasympathetic nervous activity on heart rate, the so-called "balance model". BP = Blood Pressure; HR = Heart Rate; SV = Stroke Volume (Adapted from Aubert et al. 2003).

Because of the referred link between the ANS and HRV, the variations in inter-beat intervals are commonly used as an index of autonomic responsiveness for either clinical or exercise science related purposes, representing the autonomic balance between the sympathetic and parasympathetic pathways acting on the intrinsic rhythm of the Sinoatrial node of the heart (Achten & Jeukendrup 2003; Freeman et al. 2006).

HRV analysis can be performed either by using time-domain, frequency-domain or the more recent non-linear dynamics methods. However, the frequency domain method is most frequently used for the short-term assessment of HRV, which is the case in the cardiopulmonary exercise testing. The High Frequency (HF) component of HRV obtained by frequency analysis represents parasympathetic activity upon the heart and is modified by the breathing frequency (BF). The Low Frequency component (LF) is now accepted to reflect the activity of both the Parasympathetic and Sympathetic Nervous Systems, even though it was initially thought to only be linked to the latter. (Hartikainen et al. 1998.)

HRV's measurement at rest predicts fatal arrhythmias and death due to cardiac causes, raising great interest in the field. High HRV might be a reflection of high autonomic responsiveness, and is associated with high VO_2max values. Low HRV values may be a reflection of elevated sympathetic activity that, among other things, attenuates baroreceptor reflexes. Lowered HRV may also reflect an intrinsic impairment in the physiological regulatory and adaptive mechanisms that regulate HR, or in other words, lowered autonomic responsiveness, causing the individual to be less able to tolerate a perturbation (Freeman et al. 2006; Achten & Jeukendrup 2003). Additionally, it is associated with overtraining, increased mortality, the incidence of new cardiac events and risk of sudden cardiac death in asymptomatic patients (Achten & Jeukendrup 2003).

HRV's utilization during exercise situations is a more recent aspect. This is due to the difficulty of isolating the signal from non-desired external noises when the HRV data is collected in exercise conditions. Nevertheless, attempts have been delivered with different purposes suggesting that it is a research area worthy of study. Examples are the studies related to the assessment of the training load (Buchheit et al. 2004; Kaikkonen et al. 2012), overtraining syndrome (Bosquet et al. 2003; Hynynen et al. 2006), or the VTs.

The assessment of VTs using exclusively the data derived from HRV analysis provides interesting results. If used properly, it is a reliable, accurate, non-invasive, non-laboratory dependent, and accessible index for the assessment of VTs, providing an interesting substitute for the usual ventilatory and lactate methods. This is a 21st century approach that has been developed following the study of James et al. (1989) which for the first time showed that the VT_1 could be hypothesised from BF, and especially, the study of Anosov et al. (2000), which for the first time assessed this same threshold exclusively from the data obtained from the beat-to-beat R-R interval. To this date, this has been successful using different methodologies in exercise disciplines such as cycling, running, swimming, walking and even in ski mountaineering, but, to the best of our knowledge, never in XC skiing, which leads us to the present study.

2.1.1 Modulation during Incremental Exercise

Although HRV analysis mostly depends on autonomic activity (e.g. Task Force 1996) at rest or during the night, it has been shown that short-term R-R intervals' variability was related more to breathing patterns during exercise. Because of cardiorespiratory coupling between HRV and ventilation, known as the RSA phenomenon, both the dominant frequency of HF (fHF) and the absolute HF power (HFP) are influenced by BF and tidal volume (tV). When exercise intensity increases, especially above VT_2, the rise in BF and tV leads to an increase in both fHF and in HFP. (Buchheit et al. 2007.)

The behaviour of HRV during exercise and its interpretation is also really well documented in the study of Cottin et al. (2004). It explains that during incremental exercise, there is a progressive increase in sympathetic tone, and that at exercise intensities higher than approximately 60% VO_2max, cardiac vagal control is no longer effective. This is associated with a decrease in LF and HF that can be observed in Table 1, where it can also be appreciated that at exercise intensities below VT_1 LF is predominant to HF, whereas at exercise intensities above VT_2 HF is predominant to LF (Cottin et al. 2004). Nevertheless, it can also be appreciated that both LF and HF decrease, which comes in agreement with studies where a decrease in the overall HRV was observed during incremental exercise tests (Tulppo et al. 1996; Yamamoto et al. 1992). Therefore, HF increases only relative to LF (or in normalized units to total power), a trend that cannot be explained by the links between HRV and ANS, because it is known that there is a predominance of Sympathetic Nervous System over the Parasympathetic Nervous System during exercise.

TABLE 1. Comparison of HRV spectral components, mean HR, and mean RR between work rates below and above VT_1. HF = High Frequency; HR = Heart Rate; HRV = Heart Rate Variability; VT = First Ventilatory Threshold (Cottin et al. 2004).

Spectral Components	Exercise intensity		
	Below VT (P_{-30})	Above VT (P_{+10})	*P*
Mean RR (ms)	508 ± 40	316 ± 15	<0.001
HR (beats·min^{-1})	118 ± 9	190 ± 6	<0.001
LF + HF (ms^2)	166 ± 120	2.4 ± 1.6	<0.001
LF (ms^2)	136 ± 100	0.2 ± 0.1	<0.001
HF (ms^2)	30 ± 24	2.2 ± 1.6	<0.001

The explanation to this relies on the non-neural, and therefore mechanical, mechanism that induces HRV. Previous studies in heart transplant recipients and in humans after inhibition of the cardiac autonomic control showed an unexpected persistence of HRV during exercise, which was related to the effect of breathing. In fact, it seems that some mechanical effect on the sinus node would be the cause for the increase in HF. Therefore, the modulation of HRV seems to be as follows. During heavy exercise, autonomic control of HR seems to be less effective compared with moderate exercise conditions, whereas at rest, there is a negligible cardiac control of non-neural mechanisms, representing only 1% of total HRV. This non-neural mechanism gradually increases to become prevalent with severe exercise because there were no significant difference in HF spectral energy between heart transplant recipients and controls during exercise. (Cottin et al. 2004.)

In summary, HRV is accepted to be mainly generated by neural mechanisms at rest and also during moderate exercise, whereas the residual HRV remaining during heavy exercise is due to some mechanical effect on the sinus node. The origin of the mechanical effect is a more complicated and controversial matter. A possible mechanical effect of increased BF on the sinus node has been proposed. This seems reasonable, especially as the hyperventilation of high intensities is mainly achieved by increases in BF (tV increases occur at low-to-moderate intensities). Based on this hypothesis, Cottin et al. (2004) explained the prevalence of HF compared with LF during heavy exercise conditions reflected in their results (Table 1) by a mechanical effect that might be due to a major increase in breathing rate combined with the disappearance of cardiac autonomic control. In addition to this, it is known that the mechanical effect of the limbs (cardiolocomotor component) also has a big impact on it (Blain et al. 2009; Di Michele et al. 2012).

2.2 Relevant Heart Rate Variability Analysis Methods

Once the R-R interval signal is ready, a variety of methods have been devised to quantify HRV, ranging from simple statistical descriptors to complex nonlinear mathematical algorithms that accentuate different underlying physiological mechanisms. HRV analysis can be performed either by using time-domain, frequency-domain or the more recent non-linear dynamics methods. However, the frequency domain method is most frequently used for the short-term assessment of HRV, which is the case in the cardiopulmonary exercise testing (Task Force 1996; Hartikainen et al. 1998). The use of time-domain analysis for the assessment of VTs from HRV is limited to the first threshold (Karapetian et al. 2008; Dourado et al. 2010), whereas, to the best of our knowledge, the less common non-linear dynamics methods use for this purpose is nonexistent, and therefore its explanation exceeds the boundaries of this study.

2.2.1 Time-Domain Methods

The time-domain indices are the simplest to calculate by easy statistical methods, even in short time frames, and consequently, it is also the most commonly used HRV analysis method. In a continuous electrocardiographic record, each QRS complex is detected, and it uses statistics derived from either the so-called normal-to-normal intervals (i.e. all intervals between adjacent R-R intervals resulting from sinus node depolarisations) or either from the instantaneous HR (see Figure 3). Its main limitation is the lack of discrimination between the activities of the different autonomic branches. (Aubert et al. 2003; Freeman et al. 2006; Task Force 1996.)

Examples of relevant time domain parameters regarding the assessment of VTs from HRV are:

- AverageNN (ms): Average of all normal R-R intervals.
- STDNN (ms): Standard deviation of all normal R-R intervals.
- RMSSD (ms): Root Mean Squared Successive Difference (i.e. the square root of the mean squared differences of successive normal R-R intervals).

- MSD (ms): Mean Successive Difference (i.e. the mean absolute difference between successive normal R-R intervals). (Aubert et al. 2003; Freeman et al. 2006; Task Force 1996.)

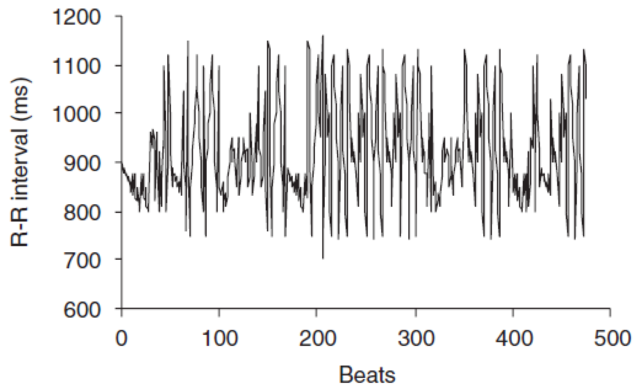

FIGURE 3. Example of an R-R interval time-series over a seven minute period at rest (Achten & Jeukendrup 2003).

2.2.2 Frequency-Domain Methods

Developments in microprocessor technology have enabled the calculation of frequency measures based on mathematical manipulations performed on the same electrocardiogram-derived data. The Frequency-Domain analysis, also named power spectra analysis or just spectral analysis, decomposes any steady, stationary, fluctuating time-dependent signal into its sinusoidal components. Therefore, instead of plotting the distribution (spectra) of HR oscillations as the change in R-R intervals over time, the power of each component is mathematically plotted as the frequency at which the length of the RR interval changes, and the computation of the power in defined frequency regions, obtaining the so-called Power Spectral Density (PSD). (Achten & Jeukendrup 2003, Aubert et al 2003 and Freeman et al 2006.)

PSD decomposes the HR signal into its frequency components and quantifies them in terms of their relative intensity, termed power, explaining how power distributes as a function of frequency. Independent of the method employed, only an estimate of the true PSD function of the signal (e.g. HR) can be obtained by proper mathematical algorithms, which would explain, in this case, how overall HRV is distributed as a function of frequency into different components. (Aubert et al. 2003; Hartikainen et al. 1998.)

The major mathematical technique for frequency analysis by which the PSD is obtained is the Fast Fourier Transform (FFT) algorithm which can plot the relative energy of different frequency components of HRV. Based on the Fourier theorem, it states that periodic signals like R-R time series can be expressed as a sum of an infinite set of sine and cosine functions with different characteristic periods of oscillation and different explanation levels of the original signal (i.e. weighting coefficients). In this way, the relative power of each point in the frequency domain can be obtained. The main reasons for the popularity of FFT analysis among the scientific community is that it is computationally efficient (simple algorithm and faster processing), gives a clear graphical representation and is readily available for application on computers. (Aubert et al. 2003; Freeman et al. 2006; Task Force 1996.)

The other main mathematical technique for frequency analysis is the Autoregressive (AR) Modelling. This approach considers the time series as a difference equation, where the signal at every time step is expressed as a linear function of its values at the order of the parametric model's previous time steps. Therefore, it requires an initial choice of the value of the order of the parametric model to provide the best fit to the data that are being processed. Even if both methods provide very comparable results, the advantages of AR are a better spectral resolution for short-term measurements, decomposition of the spectrum without the need for predefined spectral bands, and smoother and more easily interpretable shapes. This last property, the fact that AR spectrum presents smoother spectral components, is what explains its popularity in the field of the assessment of VTs from HRV, as it allows to clearly determine a cut-off between different main frequencies. (Aubert et al. 2003; Hartikainen et al. 1998.)

The PSD is usually divided into three main spectral components (Figure 4): Very Low Frequency (<0,04 Hz), Low Frequency (LF: 0.04-0.15 Hz) and High Frequency (HF: 0.15-0.4 Hz). Higher frequencies are usually termed as Very High Frequencies and are often excluded because of a poor presence and unclear physiological meaning. A similar thing happens with Very Low Frequencies, as when obtained during relatively short recordings representing a dubious measure that should be avoided when interpreting the PSD. Its possible physiological significance has been linked to thermoregulation and sympathetic discharge, but also to parasympathetic activity, as disappearance of more than 90 % of Very Low Frequency power after parasympathetic blockade suggests that HRV at all frequencies reflect mainly parasympathetic activity (Taylor et al. 1998). However, the physiological interpretations of LF and especially HF are relatively well known. The HF component of HRV represents parasympathetic activity upon the heart and is modified by the breathing. The LF is now accepted to reflect the activity of both PNS and SNS, even though in the first studies it was thought to only be linked to SNS. The HF and LF bands are usually quantified by their power values (i.e. obtained from the integral of the curve that represents the area underneath) in absolute terms (ms^2) or in relative units (n.u.) as a percentage of total power, which represent the relative value of each power component in proportion to the total power minus the VLF power component. In this study, the powers of HF and LF will be referred to as HFP and LFP, respectively (Achten & Jeukendrup 2003; Freeman et al. 2006; Hartikainen et al. 1998; Task Force 1996).

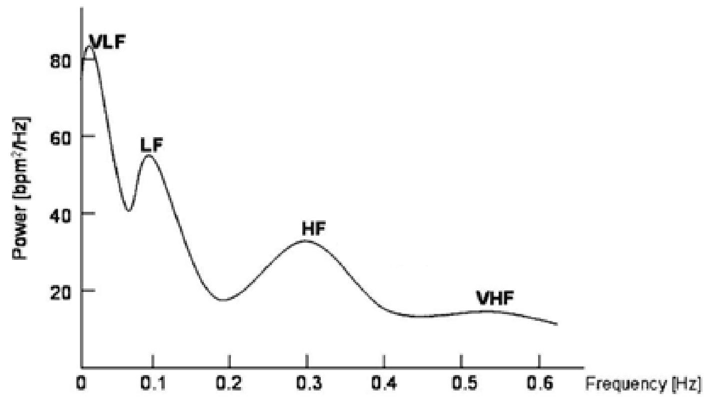

FIGURE 4. Graphical representation of a fictitious PSD and its main components obtained by a frequency domain analysis method of a RRI signal. HF = High Frequency; LH = Low Frequency; PSD = Power Spectral Density; VHF = Very High Frequency; VLF = Very Low Frequency. (Adapted from Freeman et al. 2006).

The main parameters regarding the assessment of VTs from HRV are the following:

- Low Frequency Power (LFP; ms^2): The power in the HR power spectrum between 0.04 and 0.15 Hz.
- High Frequency Power (HFP; ms^2): The power in the HR power spectrum between 0.15 and 0.4 Hz.
- Dominant Frequency on the High Frequency range (fHF; Hz): The dominant frequency in the HR power spectrum between 0.15 and 0.4 Hz. It closely corresponds to the BF.
- LFP/HFP: The ratio between LFP and HFP.
- Total Power (TP; ms^2): The power in the heart rate power spectrum at frequencies $\leq$ 0.4Hz, which represents the variances of all normal R-R intervals. (Aubert et al. 2003; Freeman et al. 2006; Task Force 1996.)

The Frequency-Domain method group, however, can be further divided into two main spectral analysis groups: non time-varying spectral analysis methods and time-varying spectral analysis methods (also referred as Time-Frequency domain methods). The first obtains a PSD of two axes (power spectral density and frequency) for a given timeframe, whereas the second obtains a tri-dimensional PSD (power spectral density, frequency and time) by making and merging several PSDs from the same time-series in a spectra vs. time called spectrogram. (Blain et al. 2005b, Cottin et al. 2006.)

The main mathematical technique of these characteristics is the Short Term Fourier Transform, which is, in a simplified way, a combination of several FFTs. It consists in choosing a small enough window for analysis in which the signal can be considered stationary in order to be able to apply classical FFT to it. The same analysis window is then applied to the rest signal blocks until the end of the processed series. The Short Term Fourier Transform is finally constituted by all the FFT performed on the successive signal blocks that are determined by regular translation of the chosen window. Another example of a time-frequency domain method is the Smoothed Pseudo Wigner-Ville Distribution, which also provides a continuous evaluation of the amplitude and frequency, giving nearly “instantaneous” complex FFT spectrum for each beat, with a high resolution achieved by independent time and frequency smoothing. (Cottin et al 2006 & 2007.)

2.3 Development of the Assessment of Ventilatory Thresholds from Heart Rate Variability

The assessment of VTs from the HRV data obtained in incremental tests is, as previously explained, a relatively new trend that is a little over a decade old. However, this fact does not exclude that, in its short period of existence, there has been substantial advances. In this section, a brief chronological overview of the main advances on this area is made to understand the development of this novel approach. Table 2 aims to guide the reader throughout the section by presenting the most relevant findings.

TABLE 2. Chronological list of the most relevant studies on the assessment of VTs from Heart Rate Variability. BF = Breathing Frequency; fHF = Dominant frequency in High Frequency; fHFm = modelled fHF; HFP = High Frequency Power; HFP_{RSA} = Respiratory Sinus Arrhythmia component of the High Frequency Power; HRV = Heart Rate Variability; LFP = Low Frequency Power; Ln = Natural logarithm; RSA = Respiratory Sinus Arrhythmia; VT_1 = First Ventilatory Threshold; VT_2 = Second Ventilatory Threshold.

Studies	Main Findings
James et al. 1989	VT_1 from BF (Cycling)
Jones & Doust 1998	Impossibility to assess VT_1 from BF (Running)
Anosov et al. 2000	VT_1 from fHF of HRV (Cycling)
Cottin et al. 2004	VT_1 from LFP:HFP of HRV (Cycling)
Blain et al. 2005b	VT_1 & VT_2 from fHF of HRV (Cycling)
Cottin et al. 2006	VT_1 & VT_2 from fHF·HFP of HRV (Cycling)
Cottin et al. 2007	VT_1 & VT_2 from fHF·HFP of HRV (Running)
Buchheit et al. 2007	VT_2 from Ln(fHFm·HFP) of HRV (Running)
Karapetian et al. 2008	VT_1 from Time Domain Analysis of HRV (Cycling)
Blain et al. 2009	Existence of a locomotor component on HFP of HRV (Cycling)
Dourado et al. 2010	VT_1 from Time Dom. Anal. of HRV, Field Test (Walking)
Di Michele et al. 2012	VT_2 from HFP_{RSA} of HRV, non time-varying (Swimming)
Mourot et al. 2014	VT_2 from Ln(fHFm·HFP_{RSA}) & HFP_{RSA} of HRV (Ski-mountaineering)

As it was just explained, the first gate to this HRV-based approach was opened by the study of James et al. (1989), which revelled that the VT_1 could be assessed from BF during an exhaustive incremental test on a cycle ergometer. However, almost a decade later, the study of Jones and Doust (1998) concluded that the breakpoint in BF does not provide a valid method for the assessment of VTs in most subjects during running exercise. This study made the following researches on the area focus on the cycle ergometer. Only two years later, the study of Anosov et al. (2000) became the first that managed to assess the VT_1 from HRV analysis. This was achieved by using the Frequency Domain analysis method, where the workload was plotted to the dominant frequency in the HF range of the HRV spectrum, which in this paper will be referred as fHF (Di Michele et al. 2012). This finding came from the understanding that the fHF, apart from being mediated by the PNS, also corresponds closely to BF. This knowledge started to arise from studies where fHF and BF were found to be strongly correlated in non-rapid-eye-movement sleep, quiet wakefulness, or during metronome-paced breathing (Anosov et al. 2000).

Meanwhile, a relationship between the withdrawal of vagal activity and the VT_1 was found in the early 90s. A few years later this, was confirmed as it was found that the vagal mediated HRV is non-existent at the ventilatory threshold level (Tulppo 1998, 47). Therefore, it seems clear that there is a link between the ANS status, HRV and the VTs.

Once a new approach to assess the VT_1 was born, different researchers followed the trend. The study of Cottin et al. (2004) used a different methodology. Instead of centring on the close relationship between fHF and BF, they focused on the LFP/HFP ratio. It was observed that at moderate exercise intensities, the LFP was predominant to the HFP, whereas in intense exercise, the HFP was prevalent to the LFP. Only a year later, Blain et al. (2005b) was the first study assessing both VTs from the analysis of R-R intervals by plotting the fHF to the workload as Anosov et al. (2000) did. Therefore, these studies based on cycle ergometer tests showed that both VTs could be assessed from the frequency analysis of HRV data. This is usually based on the close relationship between the fHF component and BF, a consequence of which the plotting of BF with exercise intensity gives two inflection points that very closely correspond to the inflection point found with ventilatory analysis.

Several years later, Cottin et al. (2006) showed that their own mathematical product of fHF multiplied by the spectral power in the HF range (HFPOW, once more using the terminology of Di Michele et al. 2012), is a more effective tool to assess the VTs than the fHF alone. They explained that the methodology of fHF vs. work rate does not work in 20% of the literature and that only a few studies have shown its reliability. They also clarified that the product of 2 indices amplified their concomitant increase occurring in correspondence with the VTs, making the visual detection of the thresholds easier, because fHF·HFP, apart from being related to BF, is also related to tV and vagal cardiovascular control. In addition, they found no significant differences between the VTs determined from ventilatory parameters and those determined from HRV data, even if the first threshold obtained though HRV sometimes tended to slightly underestimate the one obtained with the ventilatory data. (Cottin et al. 2006.)

An important note though, dependent on the fact that, to this date, all the successful studies regarding the assessment of VTs from HRV were executed in cycle ergometer tests, where the HRVTs are determined by the double non-linear increases of BF with cycling power. However, this approach is no longer valid with other exercise modes like running, where the relationship between BF and running speed during an incremental exhaustive ramp test is linear (Cottin et al. 2007), which explains why the previously mentioned study of Jones and Doust (1998) failed in the assessment of VTs from BF on running. Therefore, there are no inflection points when plotting fHF with running speed, which makes it impossible to determine the thresholds. However, the study of Cottin et al. (2007) also used the previously mentioned dfHF·HFP product, this time exporting it to running and with a successful outcome; the limitation was finally overcome. This was possible because, as it was previously explained, fHF·HFP apart from being influenced by BF was also influenced by vagal withdrawal and tV, which made it possible to observe two successive non-linear increases when plotted with workload. Therefore, this product allows determining the thresholds in exercise modalities where the relationship between BF and exercise speed is linear (i.e. running).

At about the same time, a running study with trained boys mathematically improved this same product to obtain Ln(dHFm·HFP) (Buchheit et al. 2007). To do so, a third-order model was applied to fHF for artefact correction, and the whole product was then log transformed to increase the readiness of its instantaneous changes over time. However, only the VT_2 was assessed with this method.

If all the previous studies were based on the Frequency Domain Analysis of HRV, which is the main trend in the assessment of VTs from HRV, only a few attempts towards a more simple methodology have also been made using the Time Domain Analysis of HRV. The study of Karapetian et al. (2008) was the first one in doing so in a cardiopulmonary exercise test on a stationary bike, showing a strong agreement between the VTs and the HRVTs. Afterwards, Dourado et al. (2010) determined the VT_1 from Time Domain analysis of HRV on an incremental shuttle walk test, being the first study to assess the VT_1 from HRV on a field test. In these two cases, the $HRVT_1$ lays where the HRV stops declining, showing the intensity where the parasympathetic nervous system is no longer active and where there is an increase in the activity of the sympathetic branch. The decision for the use of the time domain analysis was, together with two other minor reasons, mainly made because of its simplicity: it allowed the use of the software program compatible with the HR monitor and a conventional computer spreadsheet.

Furthermore, it is also important to note that the first studies, apart from being executed in cycle ergometers, were also performed with fixed pedalling rates. These decisions were made to avoid possible problems coming from the pedalling rate, suspicion based on the well-documented occurrence of Cardiolocomotor Coupling (LOC; Blain et al 2005a). Nevertheless, it was not until few years ago when the existence of a locomotor modulation of HRV corresponding to the pedalling rate on a graded cycling test was documented (Blain et al. 2009). This entails that HFP does not merely represent the amplitude of RSA (aRSA), as it was thought to, but is also affected by the movements of the limbs. Taking into account this consideration, Di Michele et al. (2012) removed the LOC component from the HFP band (i.e. HFP_{RSA}) before its analysis in order to exclusively analyse the component related to RSA. Even if only the VT_2 was assessed (even though it was paired to the LT_2), it was shown that it was a more accurate way than using the whole HFP component.

Finally, the very recent study of Mourot et al. (2014) made a very interesting comparison of the two latest methods for the assessment of the VT_2 from HRV (the methods of Buchheit el at. 2007 and Di Michele et al. 2012) and created two more methods by a direct combination of the previous ones (i.e. time varying $Ln(fHFm \cdot HFP_{RSA})$), and by simplifying this third one (i.e. an also time varying HFP_{RSA}). The best results were obtained with the fourth and second methods.

In summary, there have been several proposed methodologies for the assessment of VTs from HRV analysis. Initially, they were studied during cycle ergometer tests with paced pedalling rates, by plotting fHF (which corresponds to BF) to work rate. This was possible because the relationship between BF and the exercise intensity is non-linear in cycling. More recently, the assessment of VTs from HRV have also been shown to be possible in other exercise disciplines, mainly by plotting HFP-like components (related to aRSA) to workload, but, to the best of our knowledge, it has not been done in XC skiing yet. This could be due to its complex nature, which is dependent on multiple factors like LOC, RSA or the Locomotor-Respiratory Coupling (LRC). This behavioural complexity of the cardiac modulation makes it difficult to find accurate physiological explanations causing the inflexion-like behaviours of the different HRV-related components that allow the determination of the HRVTs.

2.4 Three Influence Axes to Take into Consideration

In a simplified way, we could say that cardiac modulation, apart from being mediated by the ANS, is affected by two other external factors: Respiration and Locomotion. These two factors, form together with HRV an influence triangle, resulting in three important influence axes on the assessment of VTs from HRV that are worthwhile to analyse: 1) RSA, refers to the neural interactions between respiration and heart; 2) LOC, refers to the mechanical interactions between body movements and heart; and 3) LRC, refers to the interactions between body limb movements and respiration. A graphical representation of this influence triangle can be observed in Figure 5.

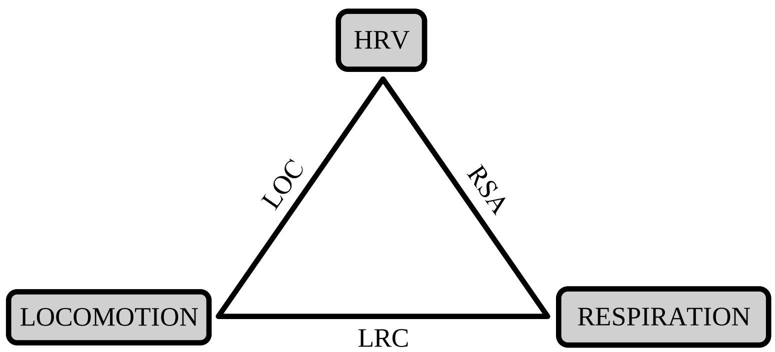

FIGURE 5. Graphical representation of the three main influence axes of Heart Rate Variability (HRV) during exercise. LOC = Cardiolocomotor Coupling; LRC = Locomotor-Respiratory Coupling; RSA = Respiratory Sinus Arrhythmia.

2.4.1 Respiratory Sinus Arrhythmia

The assessment of VTs is usually based on estimating the behaviour of ventilatory parameters from frequency-domain (sometimes also time-domain) analysis of HRV, which is possible because of RSA (Di Michele et al. 2012). RSA is a well-known cardiorespiratory coupling that results in the modulation of sinus rhythm by breathing, and as a consequence of which, there is a dominant frequency in the HF range of the HRV spectrum that corresponds to the BF. This definition can be found in the study of Blain et al. (2005a) that serves as reference in this topic.

At rest, RSA improves the efficiency of pulmonary gas exchange in matching pulmonary blood flow to lung inflation and buffers blood pressure and cardiac output variations mechanically induced by respiration. It is believed to be mainly mediated by vagal activity; however, a small degree of RSA has been observed in heart transplant subjects, meaning that human heart has a non-neural and intrinsic ability to regulate heart rate in response to respiration. This is partly why the use of the RSA power spectrum as an index of cardiac vagal tone is not widely accepted. This point of view is also supported by the fact that an increase in tV or BF can enhance or reduce RSA power, respectively, as these are not dependent on cardiac vagal control. (Blain et al. 2005a.)

Regarding RSA during exercise, from low to mild intensity its power quickly decreases in accordance with vagal withdrawal without reaching a zero value. Subsequently, RSA power has been observed to increase with the highest exercise intensities when expressed in normalized (percentage of the total power spectrum) units. This increase, which once more could not be explained by vagal activity, was attributed to mechanically induced alterations in myocardial rhythm caused by the stretching of the sinus node in response to the exercise-induced hyperventilation. This attribution was based on the assumption that, as the existence of HF variations on exercising heart transplant subject could not be explained neurally, and as BF is closely related to HF (namely RSA), it was deduced to result from a mechanical effect of BF on the sinus node. (Blain et al. 2005a.)

Therefore it seems that RSA has mainly a neural but also a mechanical mediation in rest, whereas it has a mainly mechanical but also neural mediation in intense exercise conditions. It seems that the mechanical phenomenon in the following way. During the inspiration phase, the decrease in intrathoracic pressure increases the venous return and consequently increases the filling of the right pump. This increase induces a stretch of the sinus node, provoking an increase in its activity and an increase in HR. During the expiration phase, the emptying of the lungs provokes the opposite effect with a decrease in HR. This direct, mechanically induced electrical phenomenon where the sinus node would work like a stretch receptor, is known as the Bainbridge reflex, and is reflected by an increased R-R variability and HFP in both absolute and normalized units, independent of any autonomic activity. Therefore, the dramatic increase in ventilation occurring at the VTs could also reflect thresholds in the HRV data. (Buchheit et al. 2007; Cottin et al. 2004.)

In the studies related to the assessment of VTs from HRV (Cottin et al. 2006 & 2007), HFPOW was considered to represent aRSA, which they concluded to be influenced by tV and vagal cardiovascular control as it was explained before. By that time, on a previous study of Blain et al. (2005a) it was observed that, at exercise intensities higher than 60-65% VO_2max, the aRSA is proportional to minute ventilation (VE). Consequently, because VE shows a non-linear increase when VT_2 and sometimes VT_1 is exceeded, and so does BF in a few exercise modes like cycling, a visual inspection of the HFP trend and fHF trend vs. workload allows the determinations of VTs (Cottin et al. 2006 & 2007; Di Michele et al. 2012). However, a later study of Blain et al. (2009) noted that the HFP does not merely represent the aRSA, but it also has a LOC component. These two components were successfully divided in the study of Di Michele et al. (2012), so that only the one really representing the aRSA, named as HFP_{RSA}, was used for the assessment of the VT_2.

2.4.2 Cardiolocomotor Coupling

The concept of LOC refers to the modulation exerted by the movement of the limbs on the heartbeat in dynamic cyclic exercises like running, cycling or walking. The use of this concept has been for decades reduced to what Blain et al. (2009) names as Cardiolocomotor Synchronisation. This concept refers to the synchronisation between a step cycle and R-wave occurrence, for instance, a 1:1 HR and locomotion rate synchronisation in walking and running. This synchronisation can only be observed when HR and locomotion rates are close to each other, is not modulated by workload increases, and occurs intermittently over time. It was been proposed in several studies that cardiac synchronisation optimizes blood flow to the contracting muscles and minimizes the energy cost of cardiac muscle contraction. (Blain et al 2009.)

However, LOC also includes another aspect. In the study of Blain et al. (2009) in cycle ergometer tests, it was shown that limb muscle contractions produced by cycling exercise modulates the duration of RR intervals. This resulted in a new spectral frequency component in HRV centred at the pedalling frequency, which was named Pedalling Frequency Component (PFC). It was hypothesised that the alternation of the concentric and relaxation phase during a sequence of dynamic contractions resulted in a dynamic modulation of venous return, which increased during contraction and reduced during relaxation. The increase in venous outflow in a contraction-by-contraction basis would accentuate right heart preload and activate a stretch-induced positive chronotropic response of the sinus node at the locomotor frequency. Therefore, this new element of LOC is induced mechanically by the generation of either a positive or negative chronotropic response when muscle contractions occur during systole or diastole, respectively. Moreover, this new element does not require the heart and pedalling rates to be close to each other, is consistently identified throughout a broad range of HRs, and as it increases with increase in workload, it contributes to the HRV increase observed from moderate to maximal exercise. In this study this element that was named PFC in the study of Blain et al. (2009) will be referred as LOC for detaching it from cycling, as its existence has also been found in other exercise modes such as swimming (Di Michele et al. 2012). This Italian study found evidence suggesting that the stroke action modulates the HR during front crawl swimming. Two power spectrum density peaks were observed in the HRV power spectrum, where the first peak corresponds to the RSA and the second peak corresponds to the stroke. If the synchronisation between pedalling rate and the HR was suggested to originate from a modulation of venous return by limb muscle contractions in Blain et al. (2009), it was considered more likely to be originated from some other neural or mechanical factor in swimming.

The locomotor component contributes 40 % to the overall HRV at peak exercise, and RSA contributes about 60 %. Therefore, there is a mechanically induced increase in PFC with workload increase, which is in line with previous studies indicating the existence of a non-neural mechanism at the heart level. In relation to the above, a practical implication for the assessment of VTs is that if the LOC component is not removed from HFP, the aRSA is overestimated as the LOC component significantly contributes to HFP, especially in vigorous exercise (Blain et al. 2009). Therefore, it seems that the relative increase of HFP with workload increase is caused, on one side, by the mainly mechanical actions on the heart of RSA (consequence of the breath-to-breath modulation of venous return by $\mathring{V}$E), and on the other side, by the LOC component.

2.4.3 Locomotor-Respiratory Coupling

The entrainment of breathing rhythms during rhythmic exercise has been shown in varying modes of human locomotion, including walking, running, cycling, rowing, manual wheelchair propulsion, and XC skiing (Faria 1994; Fabre et al. 2007a). Moreover, the entrainment reaches synchronisation values between breathing cycles and fore-limb movements in dogs (Crooks et al. 2012).

The study of Fabre et al. (2007a) observes that many dynamic exercises exhibit coordination of respiratory and locomotor cycles, which would imply that there are some advantages to coordination. The theoretical advantages are a possible enhancement in efficiency in the stride mechanics and a decrease of both the energy cost and the rate of perceived exertion. These advantages are hypothesised to happen because the neural drive to $\mathring{V}E$ is proportionately related to the frequency of limb movement, reflecting an excitatory influence of voluntary movement on the respiratory centre (Cottin et al. 2007). It is also known that specific training favours entrainment between stride frequency and BF in running (Cottin et al. 2007), entrainment that is also suspected to happen in XC skiing (Fabre et al. 2007a). Although there are differences in different exercise disciplines though. There is the front crawl swimming, where BF is inevitably synchronised with limb movement (Di Michele et al. 2012), and there are exercise modes such as running or XC skiing, and then we have cycling, where there are no changes in tV and BF with different pedalling frequencies (Cottin et al. 2007).

This coupling phenomenon has an impact on the assessment of VTs from HRV, because it does not allow abstraction from following the fHF trend in several exercise modalities, and only allows the use of the HFP-related components' trends as the one choice if using the frequency domain analysis method. This is explained as follows: knowing that in exercise such as the front crawl swimming or running, BF is synchronised with stride frequency, and stride frequency tends to be proportional to workload in running (Cottin et al. 2007), it can be deduced that BF is proportional to workload in those exercise disciplines. This proportionality has been proved in running (i.e. Cottin et al. 2007) and swimming (Di Michele et al. 2012), thereafter presenting no inflexion points, a phenomenon that we hypothesised to occur in XC skiing.

In the study of Cottin et al. (2007), the nature of the relationship between BF and workload and its implication on the assessment of VTs from HRV remained clear. The study explained that the decoupling between BF and pedalling frequency in cycling probably allows the non-linear increases in BF concomitant with the VTs. In running, however, a progressive increase in speed and a concomitant increase in stride frequency induces a progressive increase in BF which remains entrained by the stride frequency. Moreover, BF in running seems to be higher than in cycling. These results confirmed the different relationships between BF vs. workload during running, which is linear as we hypothesised in XC skiing, and cycling, which is non-linear. (Cottin et al. 2007.)

In XC skiing, the study of Faria (1994) in male master level XC skiers showed that during bilateral synchronous and asynchronous arm poling simulated in a Biokinetic arm ergometer, that the mean ratio of BF to Poling Frequency (PF) was 1:1.1, and therefore BF and PF were not synchronised. The study of Fabre et al. (2007a) looked at well trained XC skiers and the aim of the study was to find differences between spontaneous breathing and paced breathing (synchronisation of the expiration phase with poling time) in two different skating techniques (V2 and V2-alternate[1]). However, they found that the spontaneous and paced breathings turned out to be very similar as minimal differences were observed between their breathing pattern responses. This was speculated to be the result of an unconscious development of a tight coordination between breathing and locomotion rhythms during the subjects' extensive years of practice, as mentioned previously (Fabre et al. 2007a). From these results, it could be concluded that the LRC tends to occur in the V2 and V2-alternate skating techniques by synchronizing expiration phase with poling time and inspiration phase with upper-limb recovery time in a 1:1 ratio. The study of Fabre at al. (2007a) also found that during arm propulsion in XC skiing, primary locomotor and

[1] The V2-alternate technique uses a double poling action on every other skating leg; it is used on flat or on moderate downhills.

accessory respiratory muscles are in part the same (i.e., the muscles of the abdominal wall, the intercostals, the sternocleidomastoids, and the pectoral muscles). Because of this, in XC skiing it seems more plausible than in other exercise disciplines to achieve the desired reduction of respiratory workload and, thus, achieve the reduction of metabolic rate as a result of tight coordination between breathing and arm-propulsion cycles. Moreover, the same study (Fabre at al. 2007a) hypothesised a marked synchronisation between expiration phase and poling time could facilitate important trunk flexion and, thus, might enhance poling efficiency in XC skiing.

The studies of Holmberg et al. (2005) and Holmberg and Calbet (2007), both conducted with elite XC skiers, showed a high correlation between PFs and BFs during DP compared to DS and leg skiing (Holmberg & Calbet 2007). Moreover, in the study of Fabre et al. (2007b) where well-trained XC skiers performed submaximal tests in V2 and V2-alternate techniques, a high dependence between respiration and mechanical factors was found, and they specified that the locomotor rhythm entrains the breathing rhythm and not the other way around. This strong entrainment was explained by biomechanical factors such as the rhythmic trunk tilting with a strong tightening of the abdominal muscles along with the high activation of some of the accessory respiratory muscles during the poling motion (Fabre et al. 2007b; Holmberg et al. 2005; Holmberg & Calbet 2007).

A later study of Fabre et al. (2012), where elite male XC skiers preformed maximal treadmill tests with V2 skating technique, found that even if BF was strictly subordinate to the PF during the intense stages and presented no significant difference, during the first stages the differences were significant. This difference arose from the use of either a strictly 1:1 or 1:2 BF to PF ratio. If it was previously explained that the subordination of BF to limb movement prevents VT assessment from fHF, but the study of Fabre et al. (2012) took a step farther. They concluded that, during V2 roller-ski skating, the determination of VT_2 from ventilatory gas analysis must be used cautiously because there is too much depending on mechanical factors, as they found that the $\mathring{V}E$ was strictly entrained by the poling pattern during an incremental test. The explanation of this, resides on the fact that the mechanical and neural (i.e., increased drive from the phasic impulses from the working limbs and body position) constrains imposed on the respiratory apparatus overwhelmed the usually observed metabolic effects on $\mathring{V}E$ (i.e., increase in ventilatory drive consequent to respiratory compensation to increased CO_2 concentration caused by the metabolic acidosis). Fabre et al. (2012) also concluded, based on the previously mentioned multidisciplinary XC skiing studies (Fabre et al. 2007b; Holmberg et al. 2005; Holmberg & Calbet 2007), that this phenomenon most likely affects other exercises where the upper body is involved in the propulsion. They proposed to reduce the mechanical constrains by fixing the PF with an audio signal as a solution to this issue in a similar way that the pedalling frequency was kept constant in the first studies of the assessment of VTs from HRV made in cycle ergometers. Nevertheless, this would be too restricting for the skier, and the appropriate rhythm would be difficult to choose, as well as being too different from real on snow XC skiing. They also recommended to use the least upper-body dependent DS technique instead of the DP or skating techniques (Fabre at al. 2012).

Crooks et al. (2012) investigated whether breathing rate is synchronised with poling rate in V2 and V2-alternate skate techniques. They concluded that BF becomes more synchronised with the PF as workload increases in both techniques. This was concluded as they observed in V2-alternate, skiers synchronised one poling cycle with one breathing cycle at high efforts, whereas they poled 1.25 times per breath at low efforts, indicating that the cycles are not synchronised. In V2, skiers breathed every second poling cycle in both low and high intensities. They also found that skiers exhaled in the early stages of pole ground contact and inhaled towards the end of the ground contact and during the pole recovery phase. With these results, they concluded that chest compression during poling and chest expansion during recovery appear to be used to help the chest musculature in exhaling and inhaling, respectively, thereby saving metabolic energy for breathing at high efforts of skiing.

Finally, the study of Mourot et al. (2014) in ski-mountaineering, an exercise discipline similar to DS, concluded that the BF-PF relationship was different than a 1:1 BF to PF ratio, as they managed to differentiate the spectral power in the HF band of HRV into two components, which is only possible if the BF and PF frequencies assessed from the HF band are different.

In summary, the relationship between BF and PF seems to be as follows: during bilateral synchronous and asynchronous arm poling, BF is not entirely synchronised with PF as BF to PF ratio is 1.1:1 (Faria 1994), a not so strong entrainment that is also reflected in DS (Holmberg & Calbet 2007) or ski-mountaineering (Mourot et al. 2014). However, DP does show a very high correlation between PF and BF (Holmberg et al. 2005; Holmberg & Calbet 2007). In V2 and V2-alternate, the differences between the spontaneous and paced breathing (1:1 BF to PF ratio) patterns were minimal, indicating that both techniques use a 1:1 BF to PF ratio (Fabre et al. 2007a). However, in the later study of Fabre et al. (2012), it was found in the V2 technique subjects use either a 1:1 or 1:2 BF to PF ratio in the early stages, before switching to a 1:1 ratio at high intensities. Moreover, in the study of Crooks et al. (2012) the skiers breathed every second poling cycle (1:2 ratio) in both low and high intensities in the V2 technique. Therefore, it looks like both the 1:1 and 1:2 BF to PF ratios are used by the skiers during V2 tests in both higher and lower intensities. The study of Crooks et al. (2012) also reported that in V2-alternate skiers had a 1:1 BF to PF ratio at high efforts (instead of in every intensity as reported in the study of Fabre et al. 2007a), whereas they poled 1.25 times per breath at low efforts. They also reported that BF becomes more synchronised with the PF as workload increases in both V2 and V2-alternate techniques. Therefore, it is likely that at least the DP, V2 and V2-alternate techniques of XC skiing present a strong synchronisation between BF and PF (especially at high intensities) as speculated at the beginning of this section.

2.5 In-depth Physiological Reasons Behind the Inflection Points

As previously explained, based on Meyer et al. (2005), LT_1 corresponds to the first increase in blood lactate concentration above resting values originated from muscle glycogenolysis, and is associated to VT_1, whereas LT_2, which closely corresponds to VT_2, approximates the maximal lactate steady state representing the exercise intensity above which a continuous increase in blood lactate is unavoidable. The first rise in blood lactate concentration occurring at LT_1 leads to an over proportional increase of CO2 output as related to $\dot{V}O_2$ due to the bicarbonate buffering of the proton resulting from the dissociation of lactic acid. As a consequence of a slightly rising CO2 partial pressure, there is a compensatory increased stimulus for minute ventilation ($\dot{V}E$) mediated via the carotid bodies, and thus, $\dot{V}E$ is also increasing over proportionally (hyperventilation). As a result $\dot{V}E/\dot{V}O_2$ increases non-linearly while $\dot{V}E/\dot{V}CO2$ remains constant (Cottin et al. 2007). The workload corresponding to these events is the VT_1, termed as anaerobic threshold by Wasserman et al. (1973), as aerobic threshold by Kinderman et al. (1979) and sometimes also referred as just ventilatory threshold (Wasserman & McIlroy 1964). It is agreed to indicate the transition between moderate and intensive exercise. Moreover, the increase in sympathetic activity during intensive exercise causes an increase in the concentration of catecholamines, which may also contribute to the simultaneous increase in $\dot{V}E$ that leads to the VT_1 (Anosov et al. 2000).

VT_2 is the consequence of the inability of hyperventilation to eliminate all the CO_2, which entails an increase in $\dot{V}E/\dot{V}CO_2$ until exhaustion (Cottin et al. 2007). This is usually called the respiratory compensation point, but is also referred to as anaerobic threshold (e. g. Kindermann et al. 1979; Meyer et al. 2005). However, it is important to note that, as explained by Di Michele et al. (2012), there is a wide debate about the equivalence of the LT_2 and the VT_2, even if there is a strong trend to consider the two thresholds as being closely related (Wassermann et al. 1973; Meyer et al. 2005). As previously explained, the assessment of these two thresholds has significant practical implications. According to Dourado et al. (2010), the thresholds are fundamental to the prescription of exercises in clinical and sport settings. High intensity exercises are often contraindicated because they have been associated with greater probability of cardiovascular events and musculoskeletal injuries and lower adherence to training. Instead, moderate intensity training programs with longer duration are recommended for most middle-aged and older adults who typically present high prevalence of both sedentary lifestyle and risk factor for cardiovascular disease (Dourado et al. 2010). Meyer et al. (2005), explains the applicability of these physiological thresholds in exercise prescription and cardiopulmonary exercise testing in sedentary subjects, and also explains their applicability in competitive athletes and other healthy populations such as recreational patients. The study of Meyer et al. (2005) also gathers information from cross-sectional and longitudinal studies regarding their significance as an endurance capacity indicator, as they can reliably distinguish between different states of endurance capacity and also detect training-induced changes.

However, as it has already been explained, the assessment of the two physiological thresholds from blood lactate or ventilatory gas analysis also present some limitations, which leaded to the use of HRV for the determination of VTs. The existence of HRVTs respond to different reasons depending on the physiological mechanisms that are being followed in each methodological approach. The different methodologies could be classified in 4 groups: 1) Methods based on Time-domain analysis, 2) Method based on LFP to HFP ratio, 3) Methods based on fHF, and 4) Methods based on HFP-related components. Figure 6 presents a graphical representation of these four methodology groups.

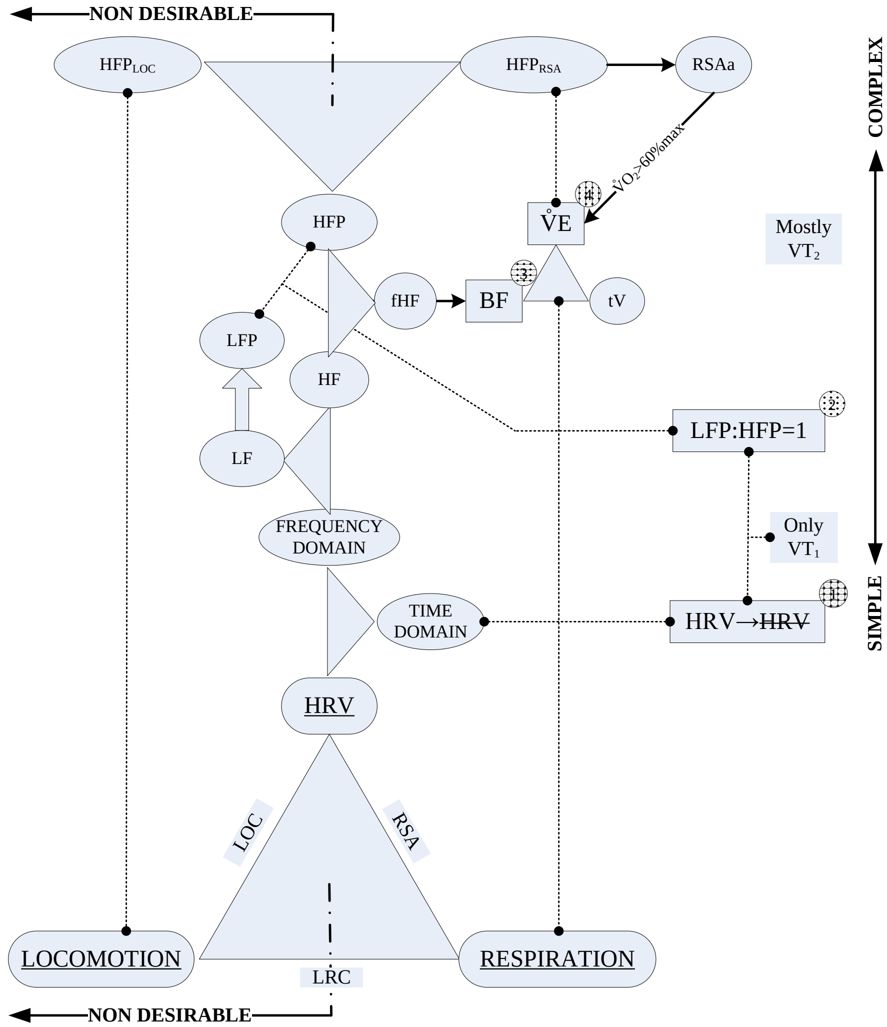

FIGURE 6. Model for understanding different methods for the assessment of Ventilatory Thresholds (VTs) from Heart Rate Variability (HRV). aRSA = Amplitude of RSA; BF = Breathing Frequency; fHF = Dominant frequency on HF range; HF & HFP = High Frequency and its power, respectively; HFP_{LOC} & HFP_{RSA} = LOC & RSA component of the HFP; LF & LFP = Low Frequency and its power, respectively; LOC = Cardiolocomotor Coupling; LRC = Locomotor-Respiratory Coupling; RSA = Respiratory Sinus Arrhythmia; tV = Tidal Volume; $\mathring{V}E$ = Minute Ventilation; VT_1 & VT_2 = First & Second VTs.

2.5.1 Methods based on Time-Domain Analysis

Time-Domain analysis has been used in the previously mentioned cycle ergometer study of Karapetian et al. (2008) and the walking study of Dourado et al. (2010). The $HRVT_1$ is found in the exercise intensity where the time-domain parameters of MSD and TDNN present no subsequent decline because HRV does not undergo any excessive variation. The reason for the progressive reduction of HRV with the increase in exercise intensity relies on the fact that vagal modulation of HR disappears almost completely at 60% of VO_2peak and thereafter the increase in HR is mainly mediated by sympathetic activation (Dourado et al. 2010). Therefore, during intensive exercise, the HRV does not undergo significant reductions and this deflection point is the result of the nearly complete disappearance of vagal modulation (i.e. Tulppo et al. 1998). This phenomenon occurs at the same relative intensity at which the lactate threshold is observed, and it indicates the moment in which parasympathetic activity decreases and sympathetic activity increases in such a way that little or no vagal regulation remains. These methods present a limitation, being that they are only valid to determine the $HRVT_1$.

2.5.2 Method based on Low Frequency to High Frequency Ratio

LF to HF ratio method was used in the previously mentioned study of Cottin et al. (2004). This method follows the same physiological explanation as the Time-Domain methods (i.e. increasing the sympathetic nervous system and decreasing the parasympathetic nervous system), but uses the frequency analysis method. In this study it was observed that, at exercise intensities higher than $HRVT_{1,}$ HFP becomes predominant to the LFP, meaning that there is a prevalence of HF in contrast to LF during heavy exercise, whereas the opposite phenomenon occurs during moderate exercise (exercise intensities lower than $HRVT_1$). However, $HRVT_1$ was never assessed as the study design consisted of two steady-state exercise bouts below and above the already known VT_1 following a warm up. The only incremental test to exhaustion was performed at the beginning of the measurements in order to know the workloads below and above the VT_1, but the HRV indexes were not reported. Nevertheless, it is possible that if the LFP and HFP would have been studied the value of the LFP/HFP would have been close to one. (Cottin et al. 2004.)

The change between the dominance of the different branches of HRV was explained as follows. In the first phases of an incremental exercise test, HFP decreases because there is a progressive withdrawal of the vagal activity to the point where at intensities over VT_1, there is no longer any cardiac vagal control or its control is no longer effective. However, after the VT_1, HFP starts increasing, even though neurally, it would be logical to be constant. Therefore, the increase in HFP needs to have a non-neural, and thus, mechanical origin, that the study of Cottin et al. (2004) explained by the mechanical effect of breathing on the sinus node. In other words, the prevalence of HFP to LFP during heavy exercise could be explained by a major increase in BF combined with the disappearance of cardiac autonomic control. Nevertheless, it is important to note that the mechanical source that causes the prevalence of HFP to LFP after $HRVT_1$ is not only a increased BF (i.e. mechanical component of RSA), but also a increased LOC component as proven in later studies. This second method also shares the same limitation as the first method group: it could only be valid to determine $HRVT_1$. (Cottin et al. 2004.)

2.5.3 Methods based on the Dominant Frequency in High Frequency Range

This method group focuses on the strong relationship between fHF and BF, and therefore, the appearance of the HRVTs is justified by the changes in BF rather than by the predominance of the different branches of the ANS. This methods have been used by the aforementioned mentioned studies of Anosov et al. (2000), Blain et al. (2005b), and Cottin et al. (2006) based on incremental tests in cycle ergometers. It is known that at low exercise levels, increasing the work intensity usually entails a marked increase in tV and a slight increase in BF, and with the increase in workload the contribution of these two components of $\mathring{V}E$ changes so that hyperventilation is mainly based on an increased BF with tV remaining constant (Cottin et al. 2006). Similar to this, the study of Blain et al. (2005a) found that at exercise intensities higher than 60-65% VO_2max, BF follows the same behaviour as $\mathring{V}E$ because the tV is already close to the maximal. Because of this, at exercise intensities of VT_1 and VT_2, the BF trend presents two non-linear increases to meet the increase in ventilatory demands. Even if these non-linear increases allow the assessment of both HRVTs, it has one limitation, it is only applicable in exercise modalities such as cycling where BF, and therefore fHF, present two non-linear increases when plotted with exercise intensity. Therefore, it is not applicable in exercise disciplines where the relationship between BF and workload is linear. The linearity between exercise intensity and BF is a consequence of the previously explained high LRC (Cottin et al. 2007).

2.5.4 Methods based on High Frequency Power related Components

This last method group is based on the close relationship between HFP and aRSA, and it allows for the assessment of both VTs in exercise disciplines where the workload-BF relationship is non-liner (like in cycling) and linear (running or swimming). aRSA is also linked to BF and vagal withdrawal, and it follows the same behaviour as $\mathring{V}E$ at exercise intensities bigger than 60 % VO_2max (Blain et al. 2005a). $\mathring{V}E$ tends to show two inflections points when plotted with workload in every exercise mode.

Moreover, it has been shown that during incremental exercise the overall HRV decreases (Cottin et al. 1999; Tulppo et al. 1996; Yamamoto et al. 1992). As a result, just before VT_1 is reached, HFP tends to be minimal (Blain et al. 2005a; Tulppo et al. 1996), but when this intensity is exceeded, recent studies have shown that HFP increases progressively (Cottin et al. 2007). This change in the HFP trend helps determine the $HRVT_1$. In relation to the assessment of the VT_2, it has been shown there is an increase in HFP when BF increases and tV remains constant in different exercise intensities. This phenomenon was caused by some non-neural mechanism, and its effect was larger at high tVs. As high tVs are achieved in high exercise intensities, BF, the related aRSA, and HFP increase more, leading to the $HRVT_2$. This increase in HFP is mainly the result of the mechanical effect of the increasing BF on the heart, but also of the increased LOC component. (Cottin et al. 2007.)

The studies of Cottin et al. (2006) in cycling and Cottin et al. (2007) in running were the first to assess VTs from HFP. In these studies the assessment of VTs from HFP werre used in combination with fHF by using the mathematical product of fHF·HFP. This product was modified to Ln(fHF·HFP) by Buchheit et al. (2007) to obtain more pronounced changes over time. All these studies used time-varying spectral analysis methods. However, Di Michele et al. (2012) used HFP on its own in a field study on swimmers, and was used with a the simpler non time-varying spectral analysis method. Therefore, instead of obtaining almost instantaneous spectral values from the PSD, a single PSD was obtained for each stage. The non time-varying spectral analysis method was used because the intermittent nature of the incremental test used in the study interrupted the steady increase in HR values, which prevents the use of a time-varying spectral analysis method. Furthermore, this study, based on the study of Blain et al. (2009), also subdivided the HFP into two components: the component truly really related to RSA (HFP_{RSA}) and the component related to the stroke or locomotor component. Even if they only assessed the VT_2, they did it using two different HFP-related components: the whole HFP component, and the component which only represented the aRSA. The results showed that assessing VT_2 by using exclusively the component related to the RSA of HFP was more accurate than using the whole component, which also represents the non-desired LOC component. The component of HFP related to RSA was more accurate because when the assessment from the whole HFP component is done, the aRSA is overestimated, an overestimation that is caused by the LOC component. In other words, HFP does not represent solely the RSA (Blain et al. 2009; Di Michele et al. 2012).

Finally, the very recent study of Mourot et al. (2014) presents an interesting comparison of four methods for the assessment of the VT_2 from HRV. The first two methods were the aforementioned methods of Buchheit et al. (2007) and Di Michele et al. (2012), respectively. The third and fourth methods were new and were based in time-domain spectral analysis methods. The third method was a direct combination of Buchheit's and Di Michele's methods, and it basically consisted on using Buchheit's method after removing the LOC-related component from the HFP component (i.e. $Ln(fHFm \cdot HFP_{RSA})$). The fourth method was a simplified form of the $Ln(fHFm \cdot HFP_{RSA})$ method, where workload was plotted to HFP_{RSA} instead of the more complex mathematical product of $Ln(fHFm \cdot HFP_{RSA})$. The best results were obtained with the fourth (i.e. HFP_{RSA}) and Di Michele's methods. (Mourot et al. 2014.)

3 RESEARCH PROBLEMS AND HYPOTHESES

The present study was conducted in an attempt to find an alternative method to assess the two submaximal thresholds other than the conventional invasive blood lactate concentration and the costly ventilatory gas exchange approaches for five different XC-skiing related techniques (i.e. NW, DP, DS, V1, and V2). The purpose of this study was to solve this main research problem by assessing these VTs from HRV analysis based on five different incremental treadmill tests to exhaustion with competitive XC skiers. Moreover, the determination of HRVTs will be made using two different HRV methods for each threshold. This task required definition of a number of hypotheses to answer the following research questions prior to the actual study:

- What is the relationship between BF and PF in the five different tests? Is there synchronisation between them?

Based on the literature, it is hypothesised that their coupling will be high in every test, and more concretely that it will be especially high in the DP test and not as high in the DS and NW tests. They would therefore present a similar BF-PF relationship compared to running where the coupling is also high (Cottin et al. 2007), and swimming, where there needs to be a perfect synchronisation (Di Michele et al. 2012). It is expected that the high coupling will result in different kinds of BF-PF synchronisation (either 1:1 and 1:2 BF to PF ratio), but also in lower levels of entrainment (e.g. 1:1.1 or 1:25 BF to PF ratio) for the different tests based on the varied findings of different studies. Nevertheless, BF-PF synchronisation will be strengthened while the workload increase (Crooks et al. 20012; Fabre et al. 2012).

We hypothesise an especially high coupling between BF and PF in the DP test maybe up to the point of presenting a 1:1 BF to PF ratio. This hypothesis was reached because DP requires a very well defined single whole-body involved locomotive action per cycle, and because the studies of Holmberg et al. (2005) and Holmberg and Calbet (2007) showed a very high correlation between them, despite the study of Faria (1994) showing a not so highly entrained 1:1.1 BF to PF ratio. Nevertheless, the study of Faria (1994) was done in a simulated bilateral synchronous poling and not in DP as the studies of Holmberg et al. (2005) and Holmberg and Calbet (2007) were. A weaker BF-PF coordination is expected for the DS test, because the importance of the upper-body movement is smaller in this discipline. Moreover, the studies from Holmberg and Calbet (2007) in DS, Mourot et al. (2014) on similar ski-mountaineering, and Faria (1994) in simulated bilateral asynchronous poling showed a not so high BF to PF entrainment. Similarly, a weaker BF-PF entrainment is also expected in NW, because this technique mimics DS and the importance of the poling in the propulsion is even more limited. Regarding the V2 test, both 1:1 and 1:2 BF to PF ratios are expected to be dominant (Fabre et al. 2007a; Crooks et al. 2012 & Fabre et al. 2012), and in the V1 test, based on the V2-alternate studies (Fabre et al. 2007a & Crooks et al. 2012), mostly a 1:1 BF to PF ratio but also a 1:1.25 ratio (especially in the lighter workloads) is expected to be found. This would mean that, in many subjects, the V1 test would show a similar BF-PF relationship as to DP.

- Which kind of relationship will the spectral peak related to RSA and the peak related to LOC present in the HF band during the five different techniques?

It is expected that the RSA component will appear at the same frequency as BF, and the LOC component will appear at the same frequency of PF. Therefore, it is expected that these two components will reflect the BF-PF relationship that was explained above. This would entail an appearance of different RSA to LOC ratios, like the 1:1 or the 1:2 ratio that was reported in the study of Di Michele et al. (2012). This seems to be very likely to occur at least in the V2 test as the studies of Crooks et al. (2012) and Fabre et al. (2012) reported a tendency for 1:2 BF to PF ratio. When BF-LF entrainment is different to the synchronised

1:1 ratio, as it would happen in the 1:2 ratio or even at entrainments of 1:1.1 ratios, two different spectral peaks should appear. The RSA peak would appear at lower frequencies than the LOC peak. Similarly, it is expected that the RSA and LOC peaks will merge into one, when there is a BF to PF ratio close to or the same as 1:1, which should happen more often with an increasing workload. This has been reported to entail a strengthened BF-PF synchronisation (Fabre et al. 2007a). Therefore, in the DP test, and perhaps in the V1 test, the two spectral peaks are expected to be almost systematically merged into one peak, which would prevent to differentiate the HF band in two spectral components. This aspect has not been reported so far as the studies dealing with RSA and LOC components (Di Michele et al. 2012 & Mourot et al. 2014) use ratios different than 1:1; for example, in the swimming study of Di Michele et al. (2012) 1:2 and 1:3 ratios were reported.

- What is the relationship between BF and workload in the five different tests? Is it linear or non-linear?

It is hypothesised that the relationship between BF and workload may be linear in the five tests, or at least close enough to linear that would prevent the appearance of two inflection points in its trend, linearity that also happens in swimming (Di Michele et al. 2012) and running (Cottin et al. 2007). This is based on the initial assumption that BF closely corresponds to PF. As stride frequency tends to increases linearly with workload in running (Cottin et al. 2007), this should also be seen in XC skiing, and therefore BF should be proportional to workload. Therefore, the assessment of VTs from HRV could not be possible based on the trend of fHF vs. workload, which is the reason why this method is discarded in the present study. Being test-specific, we hypothesise the straightest BF-workload linearity to appear in the DP test and the least linear trends in the DS and NW tests.

- Will the assessment of both VTs from ventilatory gases be successful in the five different tests?

The study of Fabre et al. (2012) concluded that, during V2 roller-ski skating, because the locomotor rhythm (PF) entrains the breathing rhythm (BF), the determination of VT_2 from ventilatory gas analysis must be used cautiously as the usual ventilatory response is altered. This was explained to occur because the neuromechanical constrain imposed on the respiratory apparatus overwhelmed the metabolic effects on $\mathring{V}E$, which complicated the detection of the inflection points used for VT_2 assessment. Based on this, it is hypothesised that the assessment of VT_2 from ventilatory gases, and most likely also the assessment of VT_1, even though there is no literature presented, will be challenging, and therefore, it might not be successful in terms of validity.

- Is the assessment of both VTs from HRV using the selected four different HRV methods possible in the five different tests?

It is expected that, in the same way as the assessment of VTs from HRV was successful in the different studies based on different exercise modes, it will also be successful in the different XC skiing tests of the present study, as minor or no changes will be introduced in the methodologies. However, the HFP_{RSA} method in the DP test is expected to be unsuccessful, because, the previously hypothesised 1:1 BF to PF ratio would entail a single merged RSA-LOC component in the HF range that would prevent the differentiation of HFP_{RSA} from HFP. This could also happen in the V1 test.

4 METHODS

4.1 Experimental Approach

A group of competitive XC skiers performed five incremental tests until exhaustion in a time frame of two to three weeks with at least one to two days between tests. The tests were performed in randomized order, anywhere from 8 a.m. to 4 p.m.

During the tests, ventilatory gases, HRV, PF, and blood lactate concentrations (blood lactate will not be analysed in this study) were collected, allowing the VTs and HRVTs to be determined. The level of agreement was then calculated between the thresholds assessed with the two methods. Volume and gas calibrations were executed twice before each test. Before warming up, the amount of subcutaneous fat was estimated based on Durnin and Rahaman (1967). Four sites skinfold measurements (m. biceps, m. triceps, os. scapula and, os. crista iliaca) were obtained on the day of the NW test. Body mass and air pressure were measured before each test, and height was measured on the first day of testing. The PF was measured with a chronometer, where ten poling cycles were timed in the last minute of each stage.

4.2 Subjects

Ten competitive national level XC skiers (five male and five female), with ages ranging from 19 to 30 years participated in the study (Table 3). All athletes were healthy and had just completed their XC ski race season in Finland. The first subjects began in the middle of April and all measurements were collected by the end of June 2013.

All subjects agreed to the terms and conditions of the study before their participation and all possible risks and benefits were explained. The study was approved by the Ethics Committee of the University of Jyväskylä.

TABLE 3. Characteristics of the study subjects (n=10).

Characteristic	**Mean ± SD**	
	Males	**Females**
Age (years)	23.4 ± 4.5	26.0 ± 3.9
Height (cm)	180.2 ± 3.1	166.6 ± 4.6
Body mass (kg)	72.2 ± 2.7	60.2 ± 4.1
Fat (%)	10.3 ± 1.5	19.4 ± 2.9
$\mathring{V}O_2$max (ml/kg/min)	73.2 ± 1.1	59.9 ± 3.5

4.3 Procedure

4.3.1 Incremental Tests

NW, DP, DS, V1 skate skiing, and V2 skate skiing were the techniques used for the five incremental tests. All the tests were performed on a large motor-driven treadmill (2.7x3.5; Rodby RL3500E, Rodby Innovations, Vänge, Sweden) located in an indoor laboratory setting in the Research Institute for Olympic Sports (KIHU, Jyväskylä, Finland). All participants wore a harness that was attached to a rope which hung from a metallic frame in the ceiling above the treadmill for safety.

Marwe Classic 800 C (rolling resistance standard / 6) and Marwe Skate 610 A (rolling resistance fast / 0) roller skis were used for all tests. Salomon and Rottefella binding systems were provided on both the skate and classic roller skis so that the athletes could use their own ski-boots. One Way poles were provided in all lengths (every 2.5 cm) with special rubber tips designed to ensure an optimal grip while roller skiing on the treadmill (Biomekanikk AS, Oslo, Norway). Athletes were allowed to select their pole length for every test and also to use their own poles, in that case the rubber tips were placed on their poles.

Before each incremental test, the subjects performed a 10 to 15 minutes warm up on the same treadmill with the same technique they were performing in the test on that day. This gave the subject a chance to warm up and feel comfortable on roller skis. The roller skis needed to be skied on 10 to 15 minutes in order to achieve the same rolling resistance throughout all tests. The incremental test was performed on the same treadmill a few minutes after the warm up, and was then followed by a 5 to 10 minutes cool down.

In all protocols, workload increased every three minutes. At the end of each three minute stage, the treadmill was stopped while blood lactate was taken from the fingertip of the subject (not used in this study). The blood lactate measurement took about 10 to 15 seconds to be obtained, while the subjects were standing. This small timeframe was included in the three minute stage.

The highest values obtained during the incremental test for 20 seconds were defined as peak values. The highest $\dot{V}O_2$peak value among the five tests was defined as the $\dot{V}O_2$max value for the person if two or more of the following criteria were met: a HR within 10 beats of the age predicted as maximal (220 – age); a respiratory exchange ratio (RER) superior to 1.1; a lactate value superior to 8.0mmol/l; or a plateau in $\dot{V}O_2$ with increasing workload. A plateau of $\dot{V}O_2$ was assumed if the $\dot{V}O_2$ of the last workload was not greater than the previous workload by 1.75 ml/kg/min (Mourot et al. 2014). A description of each protocol is in the following paragraphs.

NW Test. A combination of both inclination and speed was used to increase the workload. In this case, the protocol is best understood by looking at Table 4.

TABLE 4. Men and women's Nordic Walking tests protocol.

Workload		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Inclination	**Men**	7	9	12	14	17	20	22	23	25	27
(%)	**Women**	4	7	9	12	14	17	20	22	23	25
Speed	**Men**	6.0	6.6	6.6	7.0	7.0	7.0	7.2	7.6	7.8	7.8
(km/h)	**Women**	6.0	6.0	6.6	6.6	7.0	7.0	7.0	7.2	7.6	7.8

DP Test. Men's protocol started at a 3 % inclination and 8.0 km/h. The inclination stayed constant with speed increasing 2 km/h with every stage, up to 26km/h. Women's test started at a 3 % inclination and 7.0 km/h. The inclination stayed constant with speed increasing 2 km/h with every stage, up to 25 km/h.

DS Test. Men's test started at a 3 % inclination and 10 km/h. The inclination increased by 2 % with every stage up to 21 % while speed stayed constant. Women's test started at 3 % inclination and 9 km/h. The inclination increased by 2 % with every stage up to 21 % while speed stayed constant.

V1 Test. Men's test started at a 3 % inclination and 10 km/h. The inclination increased by 2 % with every stage up to 21 % while speed stayed constant. Women's test started at a 3 % inclination and 9 km/h. The inclination increased by 2 % with every stage up to 21 % while speed stayed constant (Figure 7).

FIGURE 7. One of the subjects during an intermediate workload of the V1 skating test.

V2 Test. Men's test started at a 5 % inclination and 8 km/h. The inclination stayed constant with speed increasing 2 km/h up to 26 km/h. Women's test started at a 5 % inclination and 7.0 km/h. The inclination stayed constant with speed increasing 2 km/h up to 25 km/h.

4.3.2 Assessment of Ventilatory Thresholds

A portable spiroergometer, the Oxycon Mobile (Viasys Healthcare GmbH, Hoechberg, Germany) was used with a proper sized mask for the gas measurements. Gas analysers and automatic volume calibrations were made twice before every test. The last calibrations were made immediately prior to the beginning of the test. The same mask size was used for each subject during every test. The main sampling unit of the spiroergometer was lightweight and attached to the upper back of the participants.

Based on Di Michele et al. (2012), the average breath-by-breath values of the last 90 seconds of every stage were considered as the subject's values for the stage, a timeframe that complies with the interval lengths considered for the HRV analysis. In the cases where the last stage was not completely finished, the average of the covered timeframe of those last 90 seconds was used as the stage value as long as a minimum of 30 seconds were completed. However, if this minimum timeframe was not completed, no value was considered as stage value, and therefore, the limited data of the last stage was excluded from the assessment of VTs. HR and VO_2 values corresponding to the VTs workloads were recorded for the use in further analyses.

VTs were determined by visual analysis of the breakpoints (Figure 8) of different variables that were well documented in the literature (Westhoff et al. 2013). The plotted variables were:

- Changes of $\mathring{V}E$ over $\mathring{V}O_2$: It presents two abrupt increases, one for each threshold.
- Changes of $\mathring{V}CO_2$ over $\mathring{V}O_2$ (i. e. V-Slope method): It presents a single abrupt increase corresponding to VT1.
- Changes of $\mathring{V}E$ over $\mathring{V}CO_2$: This graph that reflects the previously mentioned $\mathring{V}E/\mathring{V}CO_2$ is used for the determination of the VT_2, located in the main abrupt increase of its tendency.
- Changes of Ventilatory Equivalent of Carbon Dioxide ($\mathring{V}E/\mathring{V}CO_2$) and Ventilatory Equivalent of Oxygen ($\mathring{V}E/\mathring{V}O_2$) over $\mathring{V}O_2$: This graph, which allows the determination of both thresholds, was only used in the cases where the previous graphs did not allow determination of VTs. VT_1 is located at the point where $\mathring{V}E/\mathring{V}O_2$ increases non-linearly while $\mathring{V}E/\mathring{V}CO_2$ remains somewhat constant. VT_2 is located at the point where there is either the first systematic increase in the $\mathring{V}E/\mathring{V}CO_2$, or where a steeper increase in both $\mathring{V}E/\mathring{V}CO_2$ and $\mathring{V}E/\mathring{V}O_2$ is.

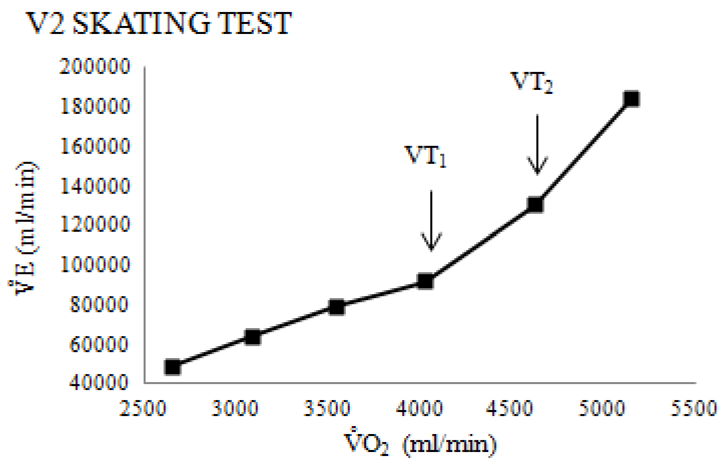

FIGURE 8. Example of the assessment of the first and second Ventilatory Thresholds (VT_1 and VT_2, respectively) from the $\mathring{V}E$-$\mathring{V}O_2$ method. $\mathring{V}E$ = Minute ventilation; $\mathring{V}O_2$ = Oxygen consumption.

4.3.3 Assessment of Heart Rate Variability Thresholds

A Suunto t6 (Suunto Oy, Vantaa, Finland) HR monitor was used with a sampling frequency of 1000 Hz for the recording of the RR intervals throughout each test. Data were downloaded to a personal computer for further analysis. From the HRV recordings, the RR interval series relative to the last 90 seconds of each step were extracted, as HF oscillations during these periods (those of interest for the study) were considered stationary.

The RR interval series were then automatically filtered by the artefact correction module from the software Firstbeat Health 3.1.1.0 (Firstbeat Technologies Oy, Jyväskylä, Finland) for the correction of the eventual artefacts and ectopic beats, which has been proven to be fast, very accurate and reliable (Saalasti et al. 2004). The automatic corrections had an affect on the beats that differed in more than 1/3 from its preceding beat and it follows either of the following rules: a) If it is considered that a single beat has been recorded divided in different fragments, these fragments are added to form a unique entity. b) If it is considered that several beats have been recorded as a unique beat, it is divided into the corresponding fragments by attributing a trend uniting the previous and forthcoming beats. An example of rule "a" and rule "b" can be observed in Table 5 A and Table 5 B, respectively. Test results where the corrected beats exceeded 5% of total heartbeats were discarded, which accounted for 10% of the total tests (five out of 50 tests).

TABLE 5. Example of automatic corrections of artefacts and ectopic beats.

Examples		R-R intervals (ms)									
A	**Recorded**	320	~~152~~	~~158~~	312						
	Corrected	320	<u>310</u>	312							
B	**Recorded**	463	~~3671~~	458							
	Corrected	463	<u>461</u>	<u>460</u>	<u>460</u>	<u>459</u>	<u>459</u>	<u>458</u>	<u>457</u>	<u>457</u>	458

The corrected R-R time series were then analysed by the software Kubios HRV 2.1 (Biosignal Analysis and Medical Imaging Group BSAMIG, Kuopio, Finland). Two different types of HRV analysis methods were used in the present study. The Time-Domain method used by Karapetian et al. (2008) was chosen for the assessment of the $HRVT_1$, whereas the non time-varying Frequency-Domain method introduced by Di Michele et al. (2012) was used for the assessment of the $HRVT_2$.

The Time-Domain parameters used in the study were MSD and STDNN. These two methods were selected because they are the most simple methods for the assessment of VT_1 found in the literature. The average values of the last 90 seconds of every stage were considered as the subject's values for the stage in order to match the interval lengths considered for the ventilatory gas analysis. In the cases where the last stage was not completely finished, the same criteria that was used with the ventilatory gases was also applied (i.e. the average of the covered timeframe of those last 90 seconds was used only if it were wider than 30 seconds; if they were not, no value was considered as stage value). The determination of the $HRVT_1$ was made by graphically plotting MSD and STDNN against workload. Next, a visual interpretation was made to locate the point where there was no further decrease in both parameters, thus indicating vagal withdrawal (Figure 9 A & 9 B). The $HRVT_1$ obtained from the MSD trend was named as $HRVT_{1\text{-}MSD}$ and the $HRVT_1$ obtained from STDNN was named as $HRVT_{1\text{-}STD}$.

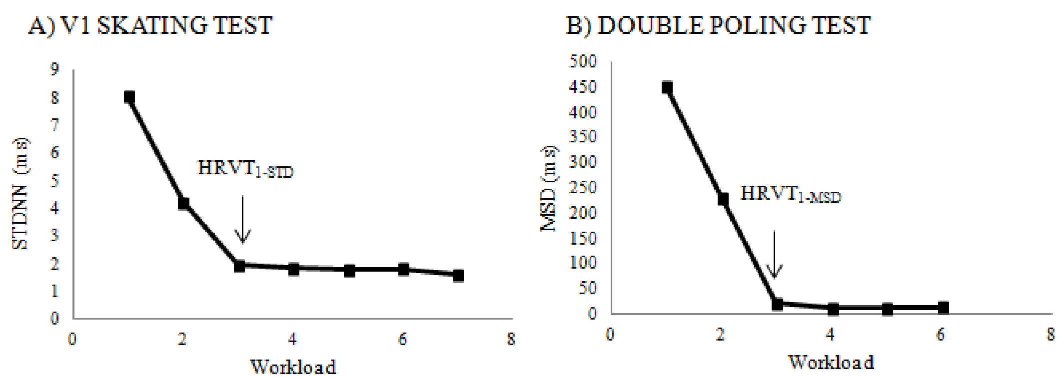

FIGURE 9. Examples of the determination of $HRVT_{1\text{-}STD}$ (9.1) and $HRVT_{1\text{-}MSD}$ (9.2) in two different incremental tests. $HRVT_{1\text{-}MSD}$ & $HRVT_{1\text{-}STD}$ = First Heart Rate Variability Threshold assessed from Mean Successive Difference (MSD) and Standard Deviation (STD), respectively.

Regarding the assessment of $HRVT_2$, two non time-varying frequency-domain analysis methods were used, which were based on HFP rather than fHF, due to the fact that, the later would be expected to present complications as it was previously explained. A non time-varying method was selected because it is the only option due to the interrupted nature of the incremental tests. At the same time, this made the analyses simpler, and when the situation allowed, the LOC components were removed from the HF spectra in an attempt to obtain more accurate data. As the RR interval time series is an irregularly time sampled series, spectrum estimation methods cannot be applied directly. Thus, the signals were converted into an equidistantly sampled form by using a piecewise cubic spline interpolation method with a sampling rate of 4 Hz. HR trend and part of the LF oscillations were removed with a 2^{nd} order polynomial model before the PSD was estimated by using an AR model of order 12 or 18, depending on the type of test. An order of 12 was applied to the DS and NW tests following the studies of Di Michele et al. (2012) and Mourot et al. (2014) in exercises with asymmetric upper body locomotion; front crawl swimming and ski mountaineering, respectively. However, an AR model order was considered to be more appropriate for the analysis of DP, V1, and V2 tests in an attempt to get more accurate data as the order 12 tended to merge different peaks. However, whether an order of 18 is preferable to an order of 12 is uncertain, because the first sometimes tends to over create tendencies (peaks), which makes the data more difficult to analyze.

The LF band was set between 0.04 and 0.15 Hz as recommended by the Task Force (1996), whereas the HF band was extended from its usual frequency range (0.15-0.4 Hz) to 0.15-2 Hz. Therefore, HFP was computed as the spectral power in the HF range (0.15–2 Hz). All the spectra (n = 321: 45 maximal tests with between 5 and 10 steps) were visually inspected, where the different exercise modality tests presented different peak patterns in the HF band, sometimes allowing to differentiate two spectral components.

In the V2 tests, two clear peaks emerged as expected from the literature. The first peak closely corresponded to BF (fHF) and the second peak to PF or the rate of arm movements, which was named as pfHF. This allowed to distinguish between two spectral powers and separate the HF band in two spectral components; the component corresponding to the cardiorespiratory modulation (HFP_{RSA}) and the component corresponding to the cardiolocomotor modulation (HFP_{LOC}) (Figure 10). The cut-off frequency was set for each spectrum at the minimum PSD value between the 2 peaks.

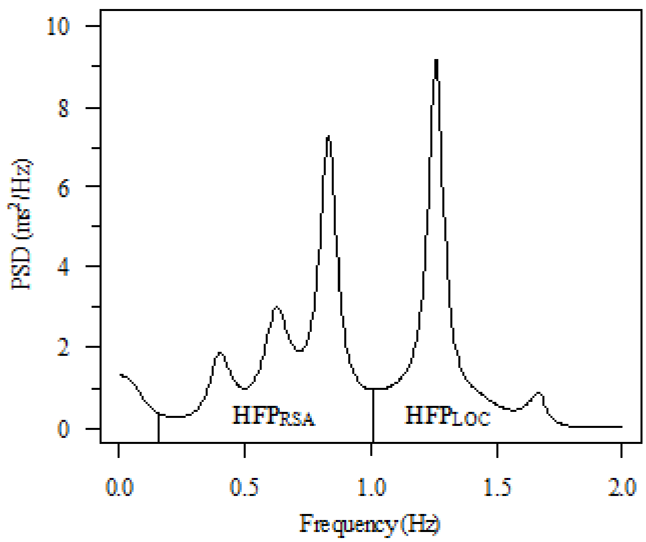

FIGURE 10. Example of the Heart Rate Variability Power Spectrum for one of the subjects in the 5th workload of the V2 test. Two Power Spectral Density (PSD) peaks are recognizable, 1 at 0.83 Hz and 1 at 1.26 Hz. The cut-off frequency separating the spectral power relative to Respiratory Sinus Arrhythmia (HFP_{RSA}) from the spectral power relative to Cardiolocomotor Coupling (HFP_{LOC}) was set at 1.01 Hz, corresponding to the minimum PSD value between the 2 peaks.

The DS (Figure 11), similar to NW but unlike the rest of the studied tests including the aforementioned V2 test, consists of a bilateral asynchronous poling pattern. As PF was measured based on whole poling cycles, each cycle consisting of two poling actions (one for each hand), means that in this case the rate of arm movements corresponds to 2·PF. In these spectras, even if the main peak corresponded to RSA, only the last and often non prominent peak's frequency seemed to be related to LOC. The frequency of this peak corresponded to the double of PF, and was therefore having a similar frequency to the rate of arm movements in front crawl swimming (Di Michele et al. 2012). Thus, in this particular test, pfHF/2 was equivalent to PF, and pfHF was not equivalent to PF. In this case, the cut-off frequency was set at the lowest PSD value between the two peaks of interest.

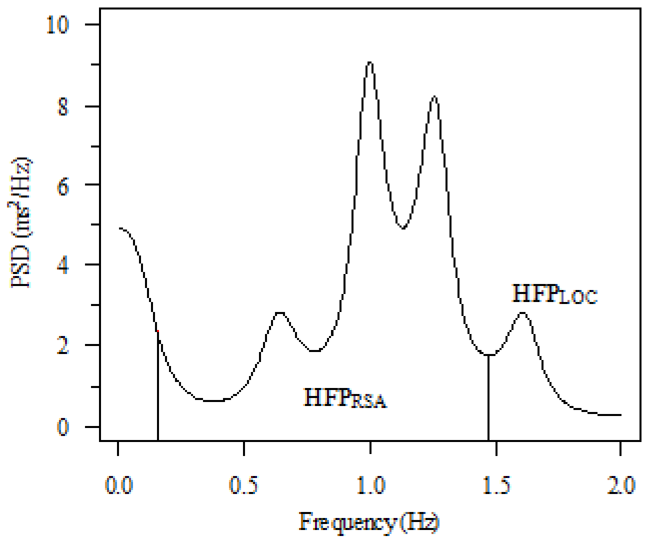

FIGURE 11. Example of the Heart Rate Variability Power Spectrum for one of the subjects in the 7th workload of the DS test. Despite several Power Spectral Density (PSD) peaks being recognizable, only the main peak at 1.00 Hz and the non-prominent peak at 1.61 Hz seem to have physiological explanation. The cut-off frequency separating the spectral power relative to Respiratory Sinus Arrhythmia (HFP_{RSA}) from the spectral power relative to Cardiolocomotor Coupling (HFP_{LOC}) was set at 1.47 Hz, corresponding to the minimum PSD value between the 2 peaks of interest.

However, unreported spectra were obtained in the NW test as their spectras presented surprisingly different patterns to the patterns obtained from the DS test, despite the fact that these two disciplines share a bilateral asynchronous poling action and therefore require a relatively similar locomotion. Even if two main peaks emerged, and despite the first peak being closely corresponded with BF, the second peak was related to PF in an unexpected manner. This peak, which was named as pfHF, closely corresponded to PF instead of 2·PF, and therefore it corresponded to the half of the rate of the arm movements, meaning that the LOC, in this case, is related to a whole poling cycle and not to each arm movement. The failure to find a peak at the double of PF was justified because the high rate of arm movements of this discipline was in most occasions higher than the upper frequency limit of PSD (2Hz) and therefore, they were not represented in PSD. The two spectral powers were divided by setting the cut-off frequency at the lowest PSD value between the 2 main peaks.

In the three aforementioned modalities a common phenomenon occurred: there was presence of a tendency for the synchronisation of PF with BF with an increase of exercise intensities. This phenomenon tended to merge the two peaks into one peak, and therefore the determination of a proper HFP_{RSA} and HFP_{LOC} was not possible. To overcome this problem, it was decided to set the cut-off frequency in the first lowest PSD value following the main peak. However this was not a solution, but an attempt to exclude the higher frequencies in order to obtain the spectrum most closely representing HFP_{RSA}. Regarding the V1 and DP (Figure 12) tests, the synchronisation between BF and PF tended to occur from the early stages, especially in DP. Thereafter this made it impossible to determine "true" HFP_{RSA} and HFP_{LOC} from the earliest stages, which is what brought us to the decision of not splitting HFP at all.

To verify that all the previously mentioned HFP_{RSA} and HFP_{LOC} components determined from the different exercise modality tests were indeed related to respiration and locomotion, the peak frequencies of these components were compared to the corresponding BF measured by the spiroergometer and the PF measured by the chronometer.

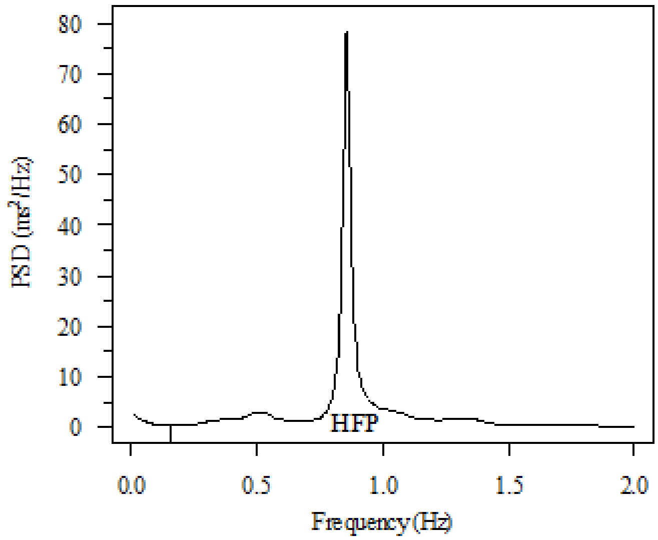

FIGURE 12. Example of the Heart Rate Variability Power Spectrum for one of the subjects in the 8th workload of the DP test. A single Power Spectral Density (PSD) peak is recognizable at 0.85 Hz, representing the frequency of both, poling and breathing rhythms. The nature of the merged peak resulted in the impossibility to set a cut-off frequency, and thus, there is a single spectral component related to both, Respiratory Sinus Arrhythmia and Cardiolocomotor Coupling (HFP).

Finally, regarding the determination of $HRVT_2$, only the full HFP and HFP_{RSA} (HFP_{LOC} was not used) of each stage were graphically plotted against the stage number. One and sometimes two increases were noticed in the trends of both variables (Figure 13 A & 13 B). A visual interpretation was made to locate the point where the discrete step from which the last increase started. The $HRVT_2$ obtained from HFP trend was named as $HRVT_{2\text{-}HFP}$ and the $HRVT_2$ obtained from STDNN was named as $HRVT_{2\text{-}HFP\text{-}RSA}$.

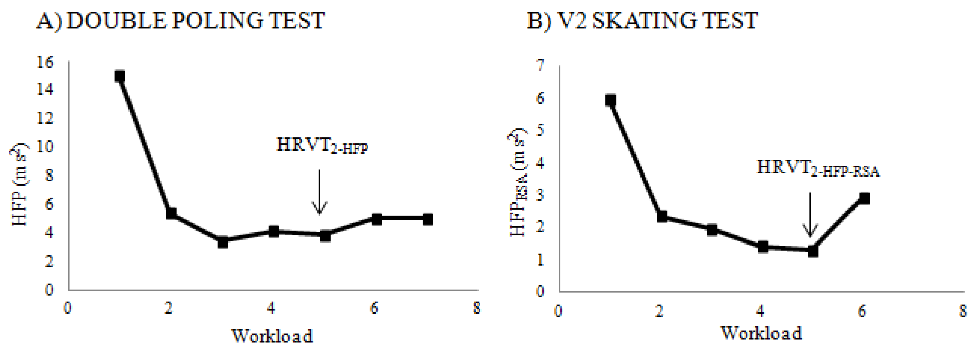

FIGURE 13. Examples of the determination of $HRVT_{2\text{-}HFP}$ (10.1) and $HRVT_{2\text{-}HFP\text{-}RSA}$ (10.2) in two different incremental tests. $HRVT_{2\text{-}HFP}$ & $HRVT_{2\text{-}HFP\text{-}RSA}$ = Second Heart Rate Variability Threshold assessed from High Frequency Power (HFP) and High Frequency Power component of Respiratory Sinus Arrhythmia (HFP_{RSA}), respectively.

4.4 Statistical Analysis

A between methods agreement was used to evaluate whether there was an agreement or bias between the variables (i.e. VTs, BF and PF) determined from the reference methods (i.e. timer for PF and ventilatory gases for the rest of variables) and the corresponding variables assessed from the alternative HRV-related methods. The analyses included (a) an evaluation of the relationships between parameters using Pearson's r correlation coefficients and linear and exponential regression lines, (b) an examination of the level of agreement using 95% limits of agreement according to Bland-Altman, and (c) a comparison of mean values using paired t tests. Even though the VTs are given in three different formats (stage Nº, % VO_2 peak and HR) the agreement analysis was only performed with HR values because of its practical applicability and to keep it simple. Apart from the agreement analysis, the mean values of PF and BF at every stage were also compared using paired t tests.

All the data are reported as the mean $\pm$ SD and the statistical significance was set at $P \leq 0.05$ for all tests. All statistical analyses were performed using IBM SPSS Statistics 20 software (SPSS Inc, Chicago, IL, USA). The normal distribution of the data was verified by One-Sample Kolmogorov-Smirnov Test (sig > 0,05). The magnitude of the correlations was assessed according to Hopkins' scale (Hopkins et al. 2009).

5 RESULTS

In every subject, the test reaching the highest $\mathring{V}O_2$peak value met the $\mathring{V}O_2$max criteria, meaning that all the subjects reached their maximal aerobic capacity in at least one of the five tests. Males' $\mathring{V}O_2$max was 73.2 ± 1.1 ml/kg/min, whereas the females' value was 59.9 ± 3.5 ml/kg/min.

5.1 Agreement between Poling and Breathing Frequencies Determined by Heart Rate Variability with their Respective Equivalents Determined Conventionally

For all the tests except for the V2 test, there was no significant difference between the mean of BF measured by HRV analysis (fHF) and the mean of BF measured by the ventilatory measurements (BF), which is the reference method. However, in the V2 test, fHF was significantly higher than BF, which was reflected in the 1.96 bias. This bias was the only one higher in absolute values than one, together with the -1.21 bias of the V1 test. The SD of the differences and, therefore, the 95 % limits of agreement were quite small for all the tests. The limits of agreement explains that 95 % of the future measurements are expected to lie within its interval. The variables were significantly correlated in all the tests, in a way that the level of correlation was very strong ($r > 0.8$) in DS, NW and V2, and was strong ($r = 0.6$-0.8) in DP and V1 (Table 6).

Regarding the agreement between PF measured by HRV analysis (pfHF) and the timed PF (PF), which is the reference method, was not as good as in BF. There were no significant differences between pfHF and PF in DP, NW and V2 tests, and the values were significantly correlated. The level of correlation was very strong in NW and V2, and was strong in DP. However, in the DS and V1 tests, pfHF was significantly higher than PF with biases of 1.91 and 1.70, respectively. Moreover, there was no significant correlation between pfHF and PF in any of these two tests ($p > 0.05$). The SD of the differences and variance intervals between pfHF and PF for all five tests were quite small as well (Table 7).

TABLE 6. Comparison between breathing frequency measured by respiratory measurements (BF) and estimated by Heart Rate Variability (HRV) analyses (fHF). r = correlation coefficient.

Test type	Agreement (breaths/min)	p	r	p	95% limits of agreement (breaths/min)
Double Poling	0.80 ± 2.80	0.445	0.718	0.045	-4.69 / +6.29
Diagonal Striding	0.18 ± 2.17	0.801	0.866	0.001	-4.07 / +4.48
Nordic Walking	0.01 ± 1.37	0.976	0.970	0.000	-2.68 / +2.70
V1 Skating	-1.21 ± 2.03	0.136	0.728	0.041	-5.19 / +2.77
V2 Skating	1.96 ± 1.92	0.016	0.912	0.001	-3.40 / +5.72

TABLE 7. Comparison between poling frequency measured by timed measurements (PF) and estimated by Heart Rate Variability (HRV) analyses (pfHF). r = correlation coefficient.

Test type	Agreement (cycles/min)	p	r	p	95% limits of agreement (cycles/min)
Double Poling	1.47 ± 2.46	0.136	0.774	0.024	-3.35 / +6.29
Diagonal Striding	1.91 ± 2.16	0.021	0.424	0.222	-2.32 / +6.14
Nordic Walking	0.65 ± 1.18	0.116	0.960	0.000	-1.66 / +2.96
V1 Skating	1.70 ± 1.76	0.029	0.541	0.166	-1.75 / +5.15
V2 Skating	0.88 ± 1.34	0.086	0.929	0.000	-1.75 / +3.51

5.2 Relationship between Breathing Frequency, Poling Frequency and Workload

5.2.1 Relationship between Breathing and Poling Frequencies with Workload

The five tests presented relatively similar tendencies regarding their mean values: there was a tendency for an increased BF-PF coupling as the workload increased and there was a higher PF than BF during the initial workloads. However three different kinds of relationships were observed after a closer look. It was also observed that there were big inter-individual differences in the type of BF-PF relationship adopted within every test.

The DP test (Figure 14 A) only showed statistically significant differences ($p \leq 0.05$) between the two parameters in the first four stages. Concerning the inter-individual differences, only one subject showed no BF-PF synchronisation at all with a BF<PF relationship during the whole test. The rest of the subjects showed either a synchronised 1:1 BF to PF ratio throughout the test (three subjects), or an initial BF<PF desynchronisation turning into a 1:1 ratio as the intensity increased (six subjects).

However, the V1 (Figure 14 D) and DS (Figure 14 B) tests showed statistically significant differences ($p \leq 0.05$) between BF and PF in more stages (stage five and six, respectively), apart from showing statistically significant differences in the initial stages, they also presented significant differences in the last stages. This happened, because the BF trend significantly surpassed the PF in both tests in a cross-like pattern, and it was also observed that the SDs were larger in the DS test than in the V1 test. Concerning the inter-individual differences, in the V1 test, most of the subjects used a clear 1:1 BF to PF ratio from the very beginning (four subjects) or they started with a BF<PF relationship to soon adopt a 1:1 ratio (three subjects). However, two subjects showed a BF-PF relationship where the initial BF<PF turned into BF>PF with the increase of intensity, in a cross-like behaviour. A single subject also showed a desynchronisation between BF and PF but without any clear pattern.

In the DS test, there was a greater inter-individual variability in the BF-PF relationship and the cross-like pattern was more predominant, as four subjects presented a clear cross-like pattern. One subject adopted a synchronised 1:1 BF to PF ratio throughout the initial workloads but later adopted a desynchronised BF>PF relationship, whereas three subjects adopted an initial desynchronised BF<PF to later adopt a synchronised 1:1 relationship. Another subject adopted a rare BF>PF relationship throughout the whole test.

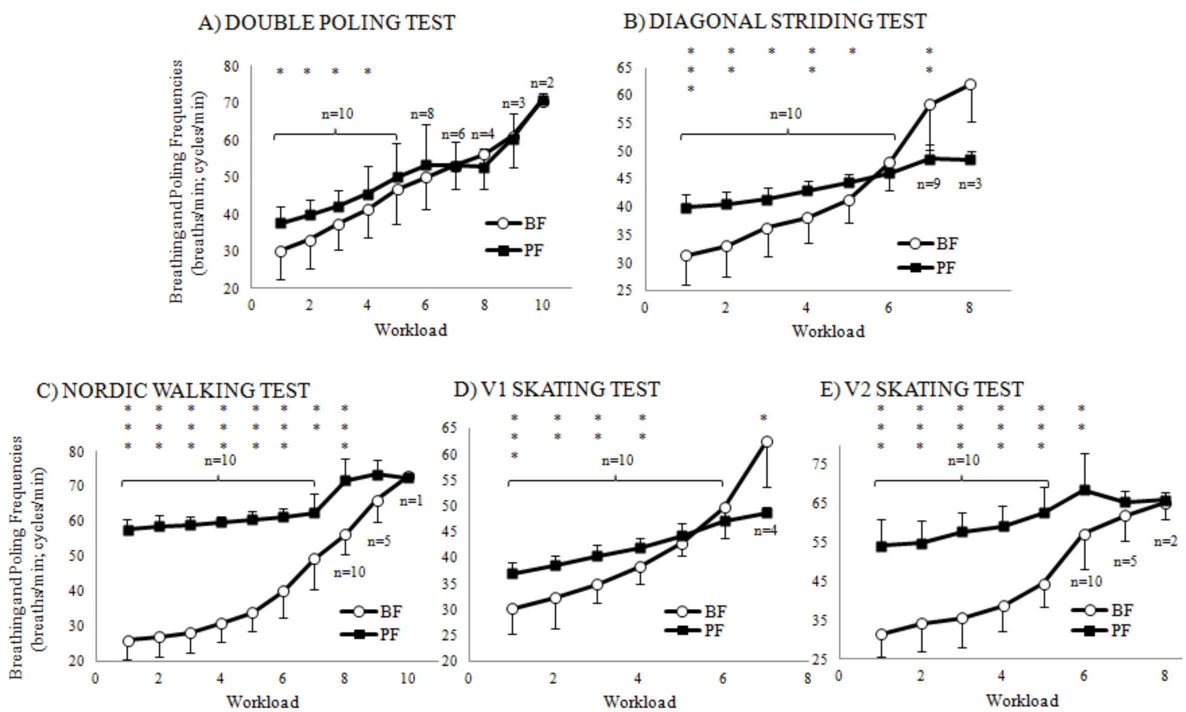

FIGURE 14. Evolution of Poling Frequency (PF) and Breathing Frequency (BF) during the five different tests. * = Significant difference between PFs and BFs, * = $p \leq 0.05$; ** = $p \leq 0.01$; *** = $p \leq 0.001$.

Regarding the NW (Figure 14 C) and V2 (Figure 14 E) tests, both tests presented a clear difference ($p \leq 0.05$) in BF and PF values except at the two highest intensities. Concerning the inter-individual differences, in the NW test, subjects adopted two different BF-PF relationship. Six subjects used a 1:2 BF to PF ratio in the early stages and later changed to a 1:1 ratio or at least got closer to it, whereas four subjects showed no BF-PF synchronisation (BF<PF) but their frequencies got closer to each other with the increase in intensity. In the V2 test, there was an even greater inter-individual variability in the adopted BF-PF relationship. One subject used a 1:1 BF to PF ratio throughout the whole test, five subjects started with 1:2 ratio and then they changed to a 1:1 ratio or at least became closer to it, and four subjects showed no BF-PF synchronisation (BF<PF) but their frequencies got closer to each other with the increase in intensity.

5.2.2 Relationship between Poling Frequency and Workload

By only focusing on the relationship of PF with workload instead of the relationship between BF and PF, the following can be observed. As seen in Figure 14, the relationship between PF and workload is quite linear for all the five tests, especially in the V1 test, even though there seems to be some change in the trend in the last stages, especially in the NW test, but also in the DP and V2 tests.

5.2.3 Relationship between Breathing Frequency and Workload

The behaviour of the mean values of BF with increasing workload intensity during the five different tests can be observed in Figure 15. It can be noticed that, even if the exponential regression line always explains a greater percentage of the total variability (R^2 range: 95.1 - 98.9 %) than what the linear regression line does, the last also explains very high amounts of the total variability (R^2 range: 89.5 - 98.3 %). The linear regression fits best to the DP test, and the V1 test is the one presenting the least fitting.

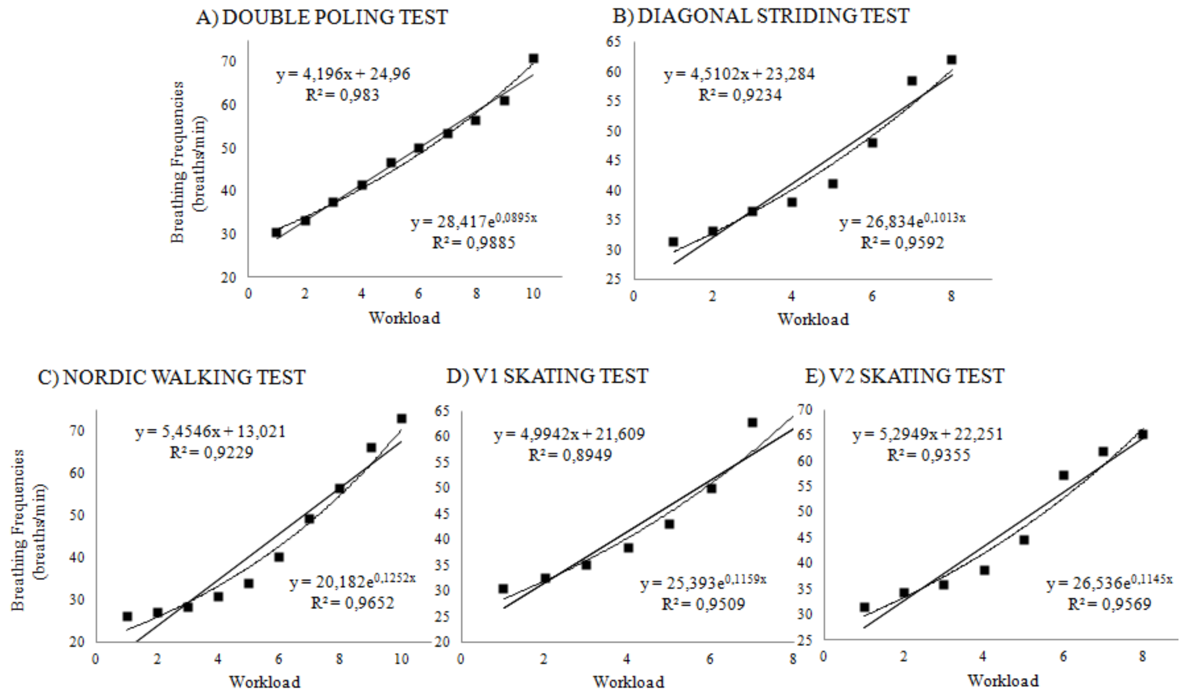

FIGURE 15. Relationship between Breathing Frequency and Workload Intensity in the five different tests. Best fitting linear and exponential regression lines are also presented together with their respective equations and R^2 values.

5.3 Assessment of Ventilatory and Heart Rate Variability Thresholds

All the VTs except the VT_1 of one subject's DP test were assessed. Regarding the determination of HRVTs, five maximal tests (2 DP tests, 2 V1 tests and 1 V2 test) were discarded because of excess of artefacts as previously mentioned. Consequently, $HRVT_1$ and $HRVT_2$ could not be assessed from any of the two methodologies. It is worth noting that three out of these five discarded tests were from the same subject. Moreover, $HRVT1_{1\text{-}MSD}$ could not be assessed in one of the DS tests because of the abnormal behaviour of its trend, and, as it was already mentioned, $HRVT_{2\text{-}HFP\text{-}RSA}$ was not assessed in DP and V1 because of the impossibility to divide the HFP spectra into two components due to a strong synchronisation between PF and BF. Other than that, the rest of HRVTs were successfully assessed, including the VT_1 that could not be assessed by the reference method. The main results regarding VT and HRVT assessment can be seen in Table 8.

TABLE 8. Main results, where the values for the first and second ventilatory thresholds determined from ventilatory gases (VT_1 and VT_2, respectively) and both Heart Rate Variability (HRV) Thresholds determined from two different HRV methods for each are expressed in different unites for the five different tests. Values are given as mean ± SD. $HRVT_{1\text{-}MSD}$ & $HRVT_{1\text{-}STD}$ = First Heart Rate Variability Threshold assessed from Mean Successive Difference and Standard Deviation, respectively; $HRVT_{2\text{-}HFP}$ & $HRVT_{2\text{-}HFP\text{-}RSA}$ = Second Heart Rate Variability Threshold assessed from High Frequency Power and High Frequency Power component of Respiratory Sinus Arrhythmia, respectively.

		n	WORKLOAD	% $\dot{V}O_2$peak	HR
DOUBLE POLING	**VT_1**	9	3.8 ± 1.0	56.9 ± 5.2	146 ± 17
	$HRVT_{1\text{-}STD}$	8	4.3 ± 0.7	69.1 ± 14.6	158 ± 10
	$HRVT_{1\text{-}MSD}$	8	2.9 ± 0.6	54.4 ± 4.9	139 ± 12
DIAGONAL STRIDING	**VT_1**	10	3.8 ± 0.8	74.1 ± 8.2	168 ± 12
	$HRVT_{1\text{-}STD}$	10	3.7 ± 0.7	73.5 ± 6.8	167 ± 12
	$HRVT_{1\text{-}MSD}$	9	3.2 ± 0.4	69.3 ± 5.7	161 ± 11
NORDIC WALKING	**VT_1**	10	4.1 ± 0.6	66.6 ± 4.4	157 ± 10
	$HRVT_{1\text{-}STD}$	10	4.8 ± 0.4	73.3 ± 5.8	166 ± 10
	$HRVT_{1\text{-}MSD}$	10	3.6 ± 0.8	62.5 ± 9.9	151 ± 15
V1 SKATING	**VT_1**	10	3.6 ± 0.7	73.5 ± 5.5	165 ± 10
	$HRVT_{1\text{-}STD}$	8	3.4 ± 0.5	71.6 ± 6.7	165 ± 8
	$HRVT_{1\text{-}MSD}$	8	2.9 ± 0.4	66.5 ± 7.2	159 ± 9
V2 SKATING	**VT_1**	10	3.5 ± 0.5	71.8 ± 6.0	165 ± 9
	$HRVT_{1\text{-}STD}$	9	3.9 ± 0.6	76.3 ± 6.6	173 ± 8
	$HRVT_{1\text{-}MSD}$	9	3.2 ± 0.7	69.6 ± 10.0	164 ± 12
DOUBLE POLING	**VT_2**	10	5.7 ± 1.7	78.6 ± 9.6	169 ± 16
	$HRVT_{2\text{-}HFP}$	8	4.6 ± 2.0	71.6 ± 15.2	160 ± 21
	$HRVT_{2\text{-}HFP\text{-}RSA}$	-	-	-	-
DIAGONAL	**VT_2**	10	5.8 ± 0.6	90.0 ± 5.9	184 ± 7

STRIDING	**$HRVT_{2\text{-}HFP}$**	10	4.7 ± 1.3	80.9 ± 12.8	173 ± 18
	$HRVT_{2\text{-}HFP\text{-}RSA}$	10	4.8 ± 1.5	82.4 ± 11.8	175 ± 17
NORDIC WALKING	**VT_2**	10	6.2 ± 0.8	86.0 ± 6.0	179 ± 9
	$HRVT_{2\text{-}HFP}$	10	5.8 ± 0.9	82.7 ± 9.2	174 ± 9
	$HRVT_{2\text{-}HFP\text{-}RSA}$	10	6.2 ± 0.8	86.0 ± 9.3	178 ± 10
V1 SKATING	**VT_2**	10	3.6 ± 0.7	73.5 ± 5.5	165 ± 10
	$HRVT_{2\text{-}HFP}$	8	3.4 ± 0.5	71.6 ± 6.7	165 ± 6
	$HRVT_{2\text{-}HFP\text{-}RSA}$	-	-	-	-
V2 SKATING	**VT_2**	10	3.5 ± 0.5	71.8 ± 6.0	165 ± 9
	$HRVT_{2\text{-}HFP}$	9	3.9 ± 0.6	76.3 ± 6.6	173 ± 8
	$HRVT_{2\text{-}HFP\text{-}RSA}$	9	3.2 ± 0.7	69.6 ± 10.0	164 ± 12

5.4 Agreement between Ventilatory and Heart Rate Variability Thresholds

5.4.1 First Threshold

DP test. Regarding the agreement between $HRVT_{1\text{-}STD}$ and VT_1, even though they were not significantly correlated ($p > 0.05$), their means were not significantly different (157 ± 10 and 147 ± 18 bpm, respectively; $p > 0.05$). However, Bland-Altman plot (Figure 16 A1) shows that $HRVT_{1\text{-}STD}$ overestimates VT_1, with a bias of 9 bpm. This plot also shows that these two variables present a very wide 95 % limit of agreement (-32 to 51 bpm) due to the wide SD of the differences (± 21 bpm). Based on the regression line (Figure 16 A2), $HRVT_{1\text{-}STD}$ only explains 0.1 % of the total variability of VT_1.

Regarding the agreement between $HRVT_{1\text{-}MSD}$ and VT_1, they were very strongly correlated ($r = 0.881$, $p = 0.009$), but the mean of $HRVT_{1\text{-}MSD}$ (138 ± 13 bpm) was significantly lower than VT_1 ($p = 0.039$) with a bias of 9 bpm. These two variables present a more narrow but still wide agreement interval (-27 to 9 bpm) (Figure 16 B1). In this case, the regression line, and therefore the equation, is much more useful, as it explains 77.6 % of the total variability (Figure 16 B2).

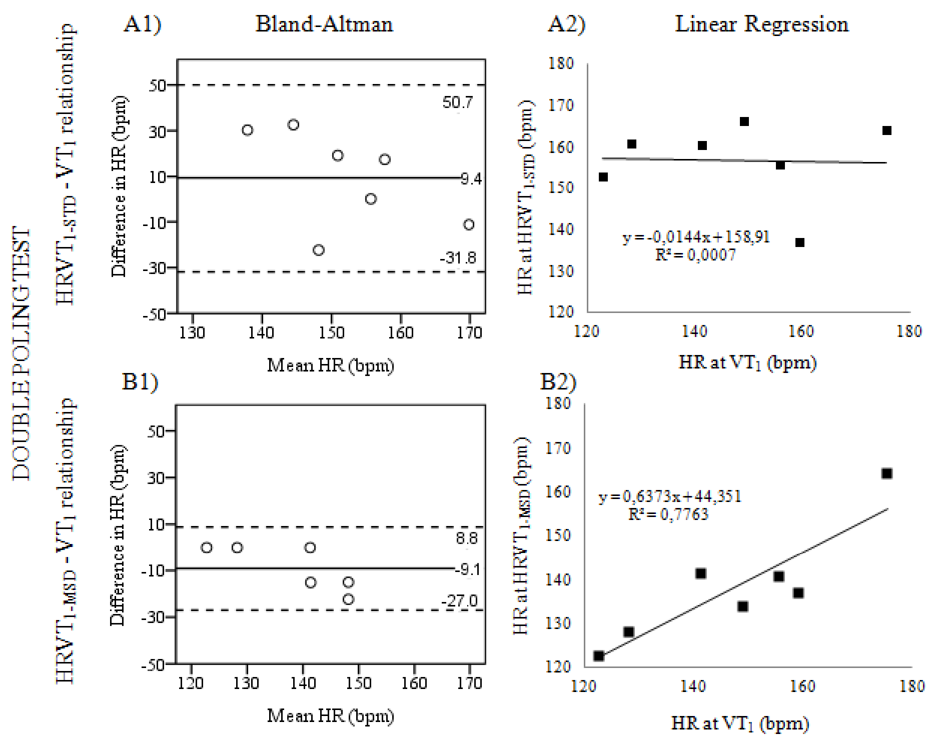

FIGURE 16. Validity testing of the assessment of the first Ventilatory Threshold (VT_1) from the first Heart Rate Variability Thresholds determined from Standard Deviation ($HRVT_{1\text{-}STD}$; Figures A1 & A2) and Mean Successive Difference ($HRVT_{1\text{-}MSD}$; Figures B1 & B2) of normal R-R intervals during Double Poling test. Bland-Altman (A1 & B1) plots the respective difference in Heart Rate (HR) between VT_1-$HRVT_{1\text{-}STD}$ and VT_1-$HRVT_{1\text{-}MSD}$ for each individual against their respective means. Dashed lines represent the limits of agreement corresponding to $\pm$ 1.96 SD. The best fitting Linear Regression lines (A2 & B2) are also displayed together with their equations, enabling a prediction of future HR values.

DS test. The best results regarding the assessment of VT_1 from HRV-derived methods were observed here. Regarding the agreement between $HRVT_{1\text{-}STD}$ and VT_1, they were strongly correlated (r = 0.772, p = 0.009), and their means were not significantly different (168 $\pm$ 12 and 167 $\pm$ 12 bpm, respectively; p > 0.05). Even if they present a relatively wide agreement interval (-17 to 15 bpm) (Figure 17 A1), $HRVT_{1\text{-}STD}$ explained 59.5 % of the total variability of VT_1 (Figure 17 A2). $HRVT_{1\text{-}MSD}$ and VT_1 were also strongly correlated (r = 0.752, p = 0.020), and their means were not significantly different either (161 $\pm$ 11 and 166 $\pm$ 12 bpm, respectively p > 0.05). The agreement interval (Figure 17 B1) and level of explanation of the regression line (Figure 17 B2) were similar to the ones between $HRVT_{1\text{-}STD}$ and VT_1.

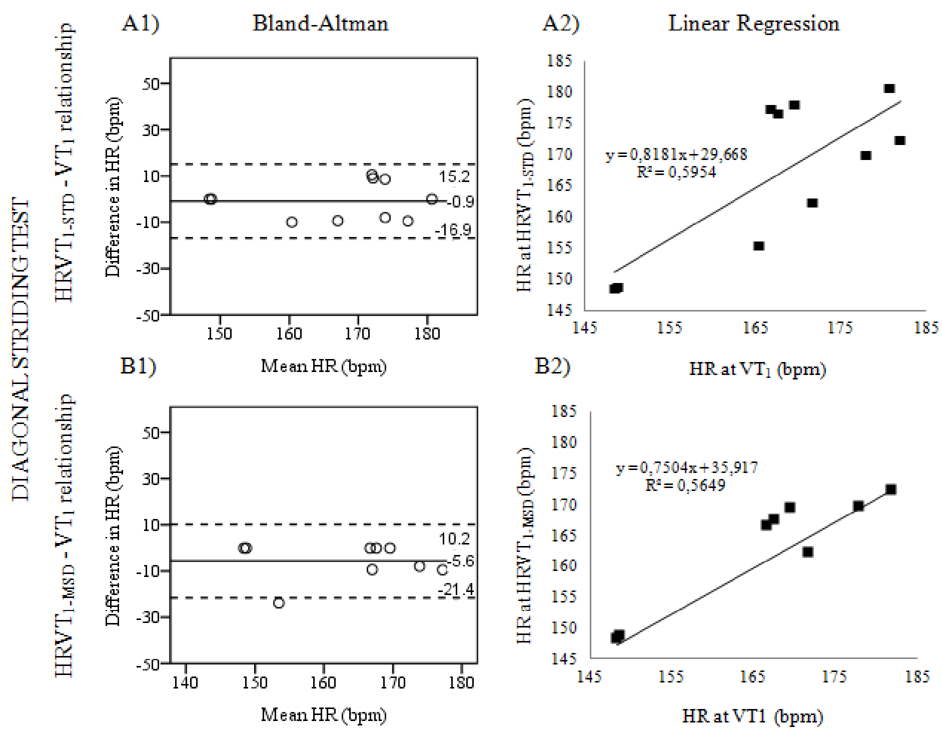

FIGURE 17. Validity testing of the assessment of VT_1 from $HRVT_{1\text{-}STD}$ (Figures A1 & A2) and $HRVT_{1\text{-}MSD}$ (Figures B1 & B2) of normal R-R intervals during Diagonal Striding test. For more details, refer to Figure 16.

NW test. $HRVT_{1\text{-}STD}$ and VT_1 were strongly correlated (r = 0.679, p = 0.031), but the mean of $HRVT_{1\text{-}STD}$ was significantly higher than VT_1 (166 $\pm$ 10 and 157 $\pm$ 10 bpm, respectively; p = 0.007), with a bias of 9 bpm (Figure 18 A1). Based on the regression line (Figure 18 A2), $HRVT_{1\text{-}STD}$ explains 46.1 % of the total variability of VT_1. Regarding the agreement between $HRVT_{1\text{-}MSD}$ and VT_1, they were not strongly correlated (p > 0.05), and the mean of $HRVT_{1\text{-}MSD}$ (151 $\pm$ 15 bpm) was not significantly different to VT_1 (p > 0.05). The Bland-Altman plot and linear regression can be seen in Figure 18 B1 and Figure 18 B2, respectively.

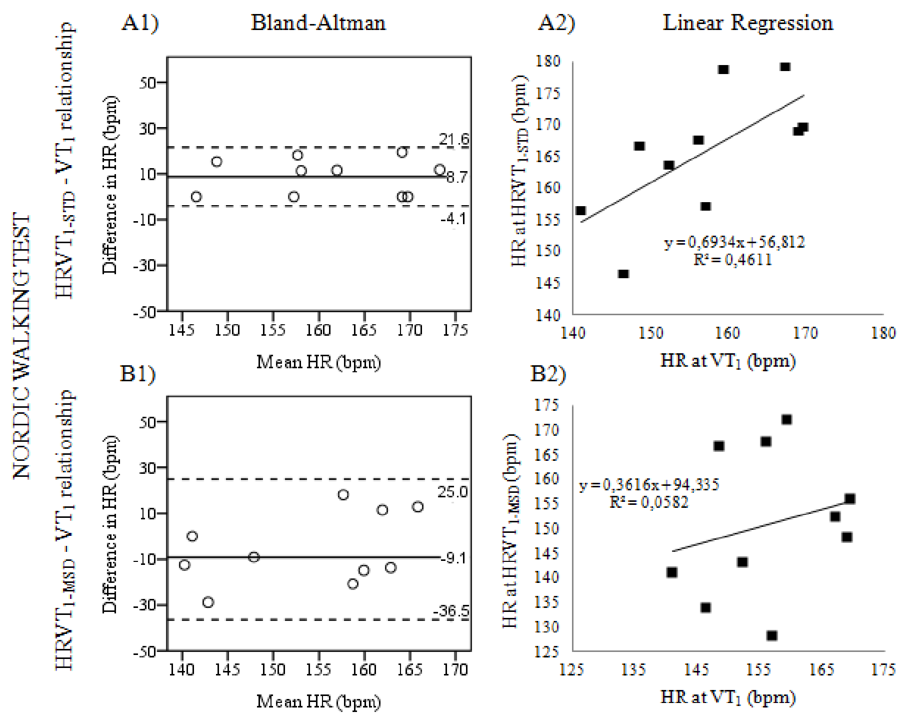

FIGURE 18. Validity testing of the assessment of VT_1 from $HRVT_{1\text{-}STD}$ (Figures A1 & A2) and $HRVT_{1\text{-}MSD}$ (Figures B1 & B2) of normal R-R intervals during Nordic Walking test. For more details, refer to Figure 16.

V1 test. Even if $HRVT_{1\text{-}STD}$ and VT_1 were not significantly correlated ($p > 0.05$), their means were not significantly different (165 ± 8 and 168 ± 8 bpm, respectively; $p > 0.05$). Indeed, the -2 bpm bias only showed a slight underestimation of $HRVT_{1\text{-}STD}$ over VT_1 (Figure 19 A1). The agreement interval (-27 to 23 bpm) was wide and the regression line (Figure 19 A2) did not show the best fit. Regarding $HRVT_{1\text{-}MSD}$ and VT_1, they were not strongly correlated either ($p > 0.05$), and, moreover, the mean of $HRVT_{1\text{-}MSD}$ (159 ± 9 bpm) was also significantly lower than VT_1 ($p = 0.049$), with a -9 bpm bias (Figure 19 B1). The regression line only explained 3.6 % of the total variability (Figure 19 B2).

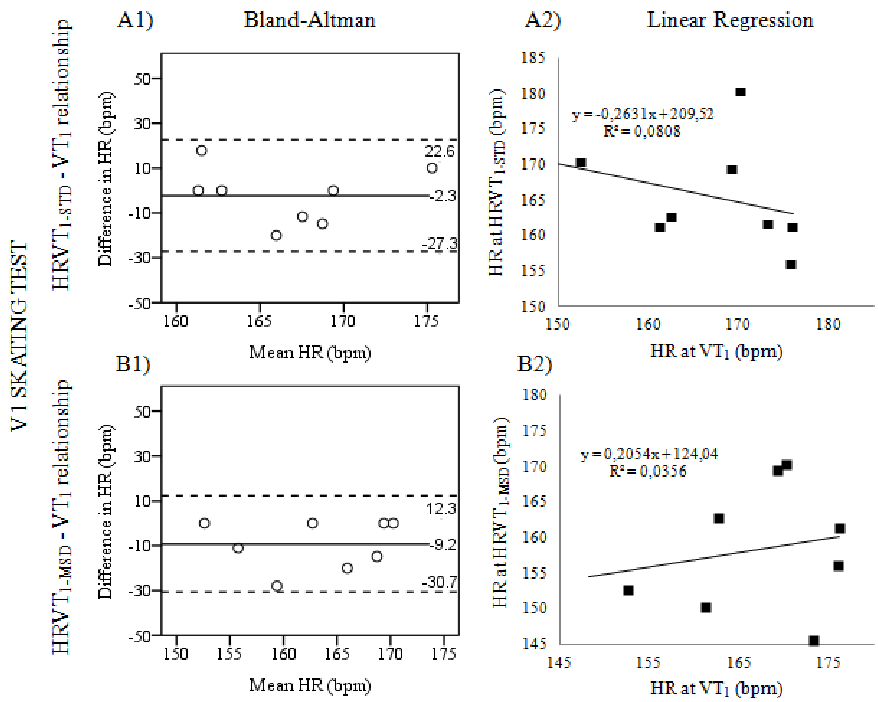

FIGURE 19. Validity testing of the assessment of VT_1 from $HRVT_{1\text{-}STD}$ (Figures A1 & A2) and $HRVT_{1\text{-}MSD}$ (Figures B1 & B2) of normal R-R intervals during V1 Skating test. For more details, refer to Figure 16.

V2 test. Despite $HRVT_{1\text{-}STD}$ and VT_1 not being significantly correlated ($p > 0.05$), their means were not significantly different (173 ± 8 and 167 ± 9 bpm, respectively; $p > 0.05$). The agreement interval was -14 to 26 bpm (Figure 20 A1) and the regression line (Figure 205 A2) explained 7.0 % of the total variability. The agreement between $HRVT_{1\text{-}MSD}$ and VT_1 was quite good. They were strongly correlated ($r = 0.765$, $p = 0.016$), and the mean of $HRVT_{1\text{-}MSD}$ (164 ± 12 bpm) was not significantly different to VT_1 ($p > 0.05$). Moreover, the agreement interval (-18 to 13 bpm) was not that large (Figure 20 B1), and the regression line explained up to 58.5 % of the total variability (Figure 20 B2).

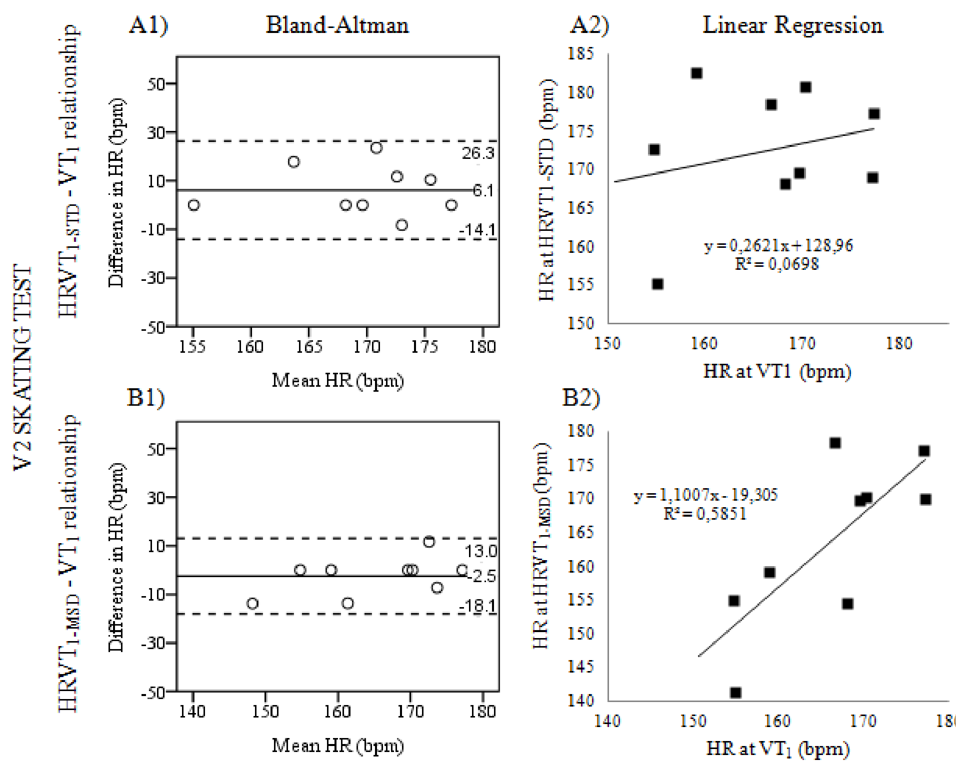

FIGURE 20. Validity testing of the assessment of VT_1 from $HRVT_{1\text{-}STD}$ (Figures A1 & A2) and $HRVT_{1\text{-}MSD}$ (Figures B1 & B2) of normal R-R intervals during V2 Skating test. For more details, refer to Figure 16.

5.4.2 Second Threshold

DP test. Regarding the agreement between $HRVT_{2\text{-}HFP}$ and VT_2, even if they were not significantly correlated ($p > 0.05$), their means were not significantly different (160.3 $\pm$ 20.6 and 168.6 $\pm$ 17.8 bpm, respectively; $p > 0.05$). However, the -8.3 bias did reflect an underestimation of $HRVT_{2\text{-}HFP}$ over VT_2 (Figure 21 A1). These two variables present a very wide agreement interval (-44.8 to 28.3 bpm) in direct consequence of the SD of the differences ($\pm$ 18.7 bpm). The regression line (Figure 21 A2) only explains 28.6 % of the total variability. As the $HRVT_{2\text{-}HFP\text{-}RSA}$ could not be assessed, there is nothing to report about the agreement between $HRVT_{2\text{-}HFP\text{-}RSA}$ and VT_2.

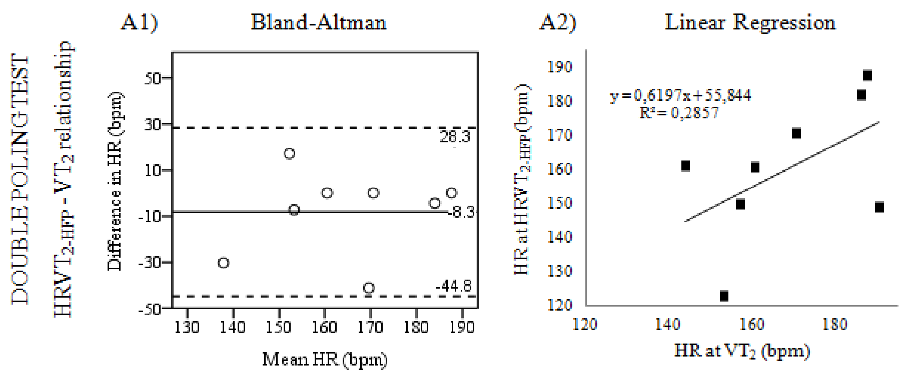

FIGURE 21. Validity testing of the assessment of the second Ventilatory Threshold (VT_2) from the second Heart Rate Variability Threshold determined from High Frequency Power ($HRVT_{2\text{-}HFP}$) during Double Poling test. Bland-Altman (A1) plots the respective difference in Heart Rate (HR) between VT_2-$HRVT_{2\text{-}HFP}$ for each individual against their mean. Dashed lines represent the limits of agreement corresponding to $\pm$ 1.96 SD. The best fitting Linear Regression line (A2) is also displayed with its equation, enabling a prediction of future HR values.

DS test. $HRVT_{2\text{-}HFP}$ and VT_2 were strongly correlated (r = 0.795, p = 0.006), but the mean of $HRVT_{2\text{-}HFP}$ was significantly lower than VT_2 (173.3 $\pm$ 17.9 and 184.0 $\pm$ 7.4 bpm, respectively; p = 0.027). This underestimation was also reflected in the bias of -10.7 bpm (Figure 22 A1). These two variables present quite a wide agreement interval (-35.8 to 14.5 bpm) and the regression line (Figure 22 A2) explains 63.1 % of the total variability. $HRVT_{2\text{-}HFP\text{-}RSA}$ and VT_2 were also strongly correlated (r = 0.664, p = 0.036), and, in this case, the mean of $HRVT_{2\text{-}HFP\text{-}RSA}$ (174.9 $\pm$ 16.8 bpm) did not differ significantly from the mean of VT_2. They present a similar agreement interval to what $HRVT_{2\text{-}HFP}$ and VT_2 present (Figure 22 B1), and the regression line explained 44.1 % of the total variability (Figure 22 B2).

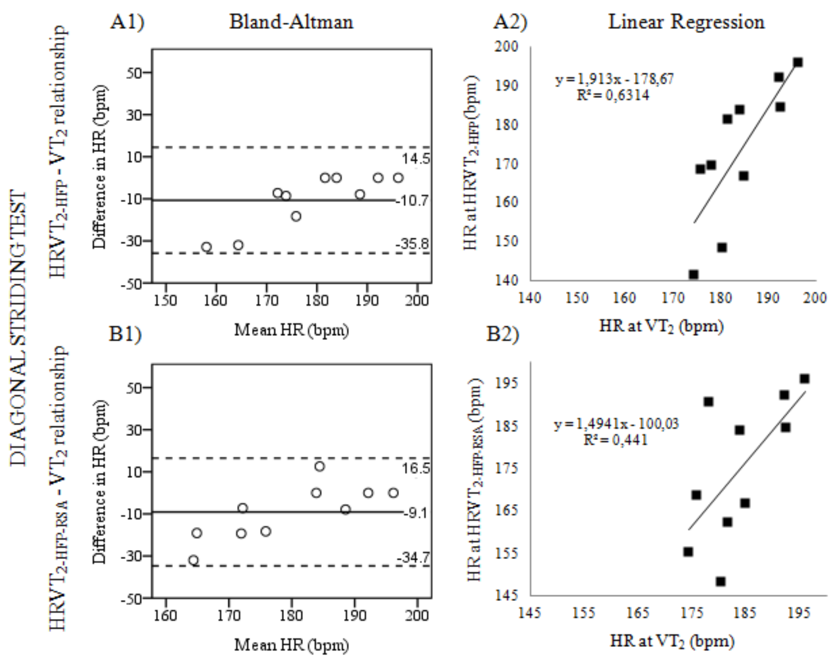

FIGURE 22. Validity testing of the assessment of the second Ventilatory Threshold (VT_2) from the second Heart Rate Variability Thresholds determined from High Frequency Power ($HRVT_{2\text{-}HFP}$; Figures A1 & A2) and High Frequency Power component related to Respiratory Sinus Arrhythmia ($HRVT_{2\text{-}HFP\text{-}RSA}$; Figures B1 & B2) during the Diagonal

Striding test. Bland-Altman (A1 & B1) plots the respective difference in Heart Rate (HR) between VT_2-$HRVT_{2\text{-}HFP}$ and VT_2-$HRVT_{2\text{-}HFP\text{-}RSA}$ for each individual against their respective means. Dashed lines represent the limits of agreement corresponding to $\pm$ 1.96 SD. The best fitting Linear Regression lines (A2 & B2) are also displayed together with their equations, enabling a prediction of future HR values.

V1 test. The agreement between $HRVT_{2\text{-}HFP}$ and VT_2 was quite poor, because these variables were not significantly correlated ($p > 0.05$) and the mean of $HRVT_{2\text{-}HFP}$ was significantly lower than VT_2 (165.4 $\pm$ 5.7 and 183.0 $\pm$ 6.5 bpm, respectively; $p = 0.001$), an underestimation that was also reflected in the -17.6 bias (Figure 23 A1). Moreover, the regression line (Figure 23 A2) only explained a poor 7.5 % of the total variability. As the $HRVT_{2\text{-}HFP\text{-}RSA}$ could not be assessed, there are no results to report about the agreement between $HRVT_{2\text{-}HFP\text{-}RSA}$ and VT_2.

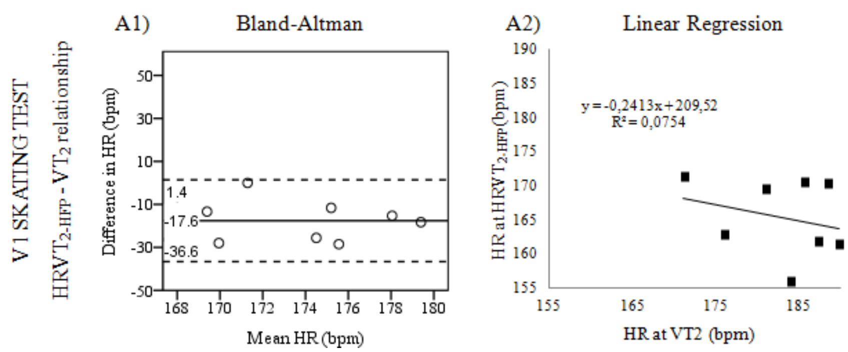

FIGURE 23. Validity testing of the assessment of VT_2 from $HRVT_{2\text{-}HFP}$ (Figures A1 & A2) during the V1 Skating test. For more details, refer to Figure 21.

NW test. Regarding the agreement between $HRVT_{2\text{-}HFP}$ and VT_2, despite not being significantly correlated ($p > 0.05$), no significant differences between the mean of $HRVT_{2\text{-}HFP}$ and VT_2 were observed (174.3 $\pm$ 8.6 and 178.6 $\pm$ 9.0 bpm, respectively; $p > 0.05$). The agreement interval was -22.6 to 14.0 bpm (Figure 24 A1) and the regression line (Figure 23 A2) only explained 19.1 % of the total variability. However, the agreement between $HRVT_{2\text{-HFP-RSA}}$ and VT_2 was close to optimal. Apart from being very strongly correlated (r = 0.818, p = 0.026) and $HRVT_{2\text{-HFP-RSA}}$ presenting no significantly different mean (178.1 $\pm$ 9.7 bpm) to VT_2 ($p > 0.05$), the bias was very close to zero (-0.5 bpm) (Figure 24 B1). Furthermore, the agreement interval (-14.9 to 13.8) was relatively small, and the regression line explained 48.31 % of the total variability as it can be appreciated in Figure 23 B2.

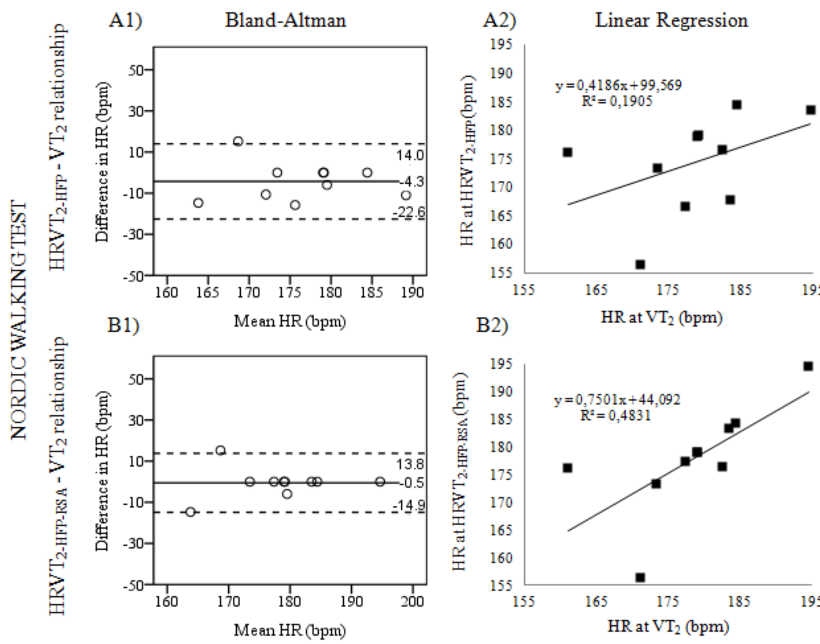

FIGURE 24. Validity testing of the assessment of VT_2 from $HRVT_{2\text{-HFP}}$ (Figures A1 & A2) and $HRVT_{2\text{-HFP-RSA}}$ (Figures B1 & B2) during the Nordic Walking test. For more details, refer to Figure 22.

V2 test. $HRVT_{2\text{-HFP}}$ and VT_2 were very strongly correlated (r = 0.809, p = 0.008), but the mean of $HRVT_{2\text{-HFP}}$ was significantly lower than VT_2 (176.5 $\pm$ 12.2 and 184.5 $\pm$ 6.0 bpm, respectively; p = 0.019). This underestimation of $HRVT_{2\text{-}HFP}$ over VT_2 was also reflected in the bias of -8.0 bpm (Figure 25 A1). They present a relatively narrow agreement interval (-24.0 to 8.0) and the regression line (Figure 25 A2) explained 65.5 % of the total variability. $HRVT_{2\text{-}HFP\text{-}RSA}$ and VT_2 were not significantly correlated ($p > 0.05$), but the mean of $HRVT_{2\text{-}HFP\text{-}RSA}$ (179.0 $\pm$ 7.4 bpm) did not differ significantly from the mean of VT_2 either. Their agreement interval was also relatively small (Figure 25 B1), but the regression line only explained a small percentage of the total variability (Figure 25 B2).

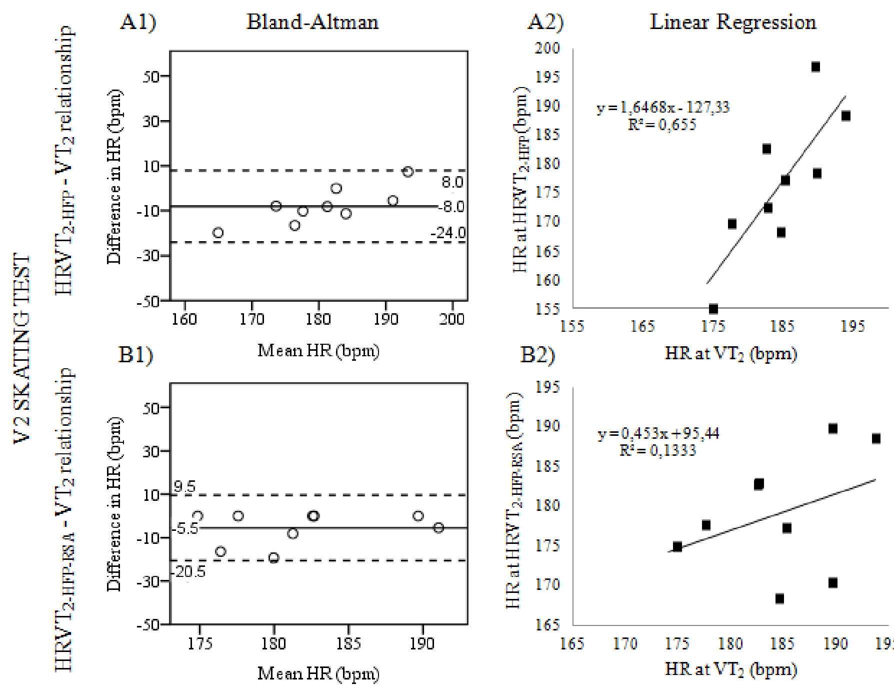

FIGURE 25. Validity testing of the assessment of VT_2 from $HRVT_{2\text{-}HFP}$ (Figures A1 & A2) and $HRVT_{2\text{-}HFP\text{-}RSA}$ (Figures B1 & B2) during the V2 Skating test. For more details, refer to Figure 22.

6 DISCUSSION

6.1 Agreement between Poling and Breathing Frequencies Determined by Heart Rate Variability with their Respective Equivalents Determined Conventionally

The BF measured by HRV analysis (fHF) and the mean of BF measured by ventilatory measurements (BF) offered almost equivalent values. Only in the V2 test did a slight but significant overestimation of fHF over BF occur. In the rest of the tests, besides the V1 test, there was also a slight but insignificant overestimation of BF when measured as fHF.

However, the agreement between PF measured by HRV analysis (pfHF) and the timed PF (PF), although good in the DP, NW, and V2 tests, was not as good in the DS and V1 tests. In these last tests, pfHF statistically overestimated PF to a slight degree, and there was no significant correlation between them. In the DP, NW and V2 tests, there was also a slight but insignificant overestimation of PF when measured by pfHF, resulting in a tendency for the overestimation of PF when measured from HRV. The smaller PF-pfHF agreement compared with BF-fHF is probably due to the fact that the spiroergometer measures the target variables in a valid way, whereas the timer is subject to considerable human error. In any case, there seems to be only a very weak tendency for the overestimation of both BF and PF when measured with HRV that does not seem strong enough to prevent the confirmation that PF and BF assessed from their respective two methodologies are indeed equivalent, or at least, that they closely correspond to each other.

Nevertheless, it is important to recall that PF values correspond to pfHF in all the tests except in the DS test, where PF corresponds to half of pfHF. This is because in the DS technique, as well as in the NW technique, there is a bilateral asynchronous poling, and thus, the rate of arm movements is $2 \cdot PF$ instead of being just PF. However, a previously unreported phenomenon occurring in NW relates the HFP_{LOC} to half of the rate of the arm movements instead of relating it to the its full rate. For this reason, in this test, the pfHF value corresponds to the frequency of the peak of the spectral component called HFP_{LOC}. Therefore, it is possible to conclude that the peaks of the spectral components HFP_{RSA} and HFP_{LOC} emerged at the expected frequencies in the PSD plots because they closely corresponded to the frequencies of BF and the rate of arm movements respectively, with the exception of the NW test, where the peak of HFP_{LOC} corresponded to the half of the rate of arm movements. This matter will be further discussed in Chapter 6.3.

6.2 Relationship between Breathing Frequency, Poling Frequency and Workload

The relevance of the relationship between BF, PF and workload in the assessment of both ventilatory and HRV thresholds, relies on the fact that a synchronisation between the first two, regardless of the type (1:1 or 1:2 BF to BF ratio), would probably bring about a linearity between BF and workload. This would worsen the assessment of both VTs and HRVTs. This was hypothesised to happen, because it was expected that PF would be proportional to workload in XC skiing in the same way that stride frequency tends to be proportional to workload in running (Cottin et al. 2007). This would entail that the same trend would be obtained when plotting workload, BF, and PF with every possible combinations, because if BF were proportional to PF, and PF were proportional to workload, then BF would also be proportional to workload. Therefore, if our hypotheses were true, the assessment of inflection points in the BF-workload or in the fHF-workload trend would not be possible. Moreover, we also hypothesised that the determination of VTs or at least VT_2, would also be complicated by the subordination of BF to PF. This was reported in the study of Fabre et al. (2012) during V2 skating. It was explained that the aeromechanical constrain imposed on the respiratory apparatus overwhelmed the metabolic effects on $\mathring{V}E$, which complicates the detection of the inflection points used for the VT_2 assessment (see Chapters 2.4.3 and 3).

As described in the results section, three different kinds of BF-PF relationships were observed throughout the five tests regarding the mean values. Roughly speaking, in the DP test these two parameters were synchronised (1:1 BF to PF ratio) almost from the very beginning. In the V1 and DS tests they showed previously undocumented statistical differences in early and late stages as the BF trend surpassed the PF trend (cross-like pattern). Finally, in the NW and V2 tests, they presented clearly different BF and PF values except at the two highest intensities. However, all the tests presented the following phenomena: a tendency for an increased BF-PF coupling with workload increase, higher PF than BF during the initial workloads, and the existence of inter-individual differences of a larger extent in the adopted BF-PF pattern. The common tendency for an increased BF-PF coupling with workload increase was not a surprise, as the study of Crooks et al. (2012) reported that BF becomes more synchronised with PF as workload increases in the two skating techniques (V2 and V2-alternate) used in the study.

In regards to the BF-PF relationship in the V2 test, the technique that has been studied the most, our results appear to be congruent with the literature. In the study of Fabre et al. (2007) the skiers used a 1:1 BF to PF ratio. However, in a more recent study of Fabre et al. (2012), it was found that skiers used either 1:1 or 1:2 BF to PF ratio in the early stages, before using a 1:1 ratio at the high intensities. Moreover, in the study of Crooks et al. (2012) the skiers breathed every second poling cycle (1:2 ratio) in both low and high intensities. Therefore, the literature shows that both, the 1:1 and 1:2 BF to PF ratios are used by the skiers during the V2 tests in both high and low intensities, which is what was partly shown by our results (i.e. these two ratios were predominant). However, some subjects presented no synchronisation between BF and PF, a phenomenon that was not reported in the aforementioned literature. Therefore, our initial hypothesis stating that there would be a high correlation between BF and PF has just been confirmed for most of the subjects when performing the V2 test, which would mean that the assessment of HRVTs using the tendency of fHF would probably be unsuccessful.

In regards to the DP test, based on the studies of Holmberg et al. (2005) and Holmberg and Calbet (2007), where a very high correlation between PF and BF was found, we hypothesised that there would be a synchronised 1:1 BF to PF ratio. The results of the present study showed that a synchronised scenario seems to be predominant in most of the subjects, especially as the workload increases. It was observed that, a single subject showed a BF<PF relationship during all the stages and most of the subjects (six out of 10) also showed a BF<PF desynchronisation during the first stages. However, these subjects soon turned into a 1:1 BF to PF ratio as the intensity increased, and three other skiers showed a synchronised 1:1 BF to PF ratio throughout the test. Therefore, our initial hypothesis stating that there would be an especially strong entrainment between BF and PF in the DP test has been confirmed to a certain extent, which would mean that the assessment of HRVTs using the tendency of fHF would probably be the most unsuccessful in this test.

In regards to the V1 test results, we are able to compare them to the studies on V2-alternate technique because of their biomechanical similarities. In the study of Fabre et al. (2007) they reported that the skiers used a 1:1 BF to PF ratio throughout the entire incremental test, whereas the study of Crooks et al. (2012) reported that skiers only had 1:1 PF to BF ratio at high efforts. In Crooks et al (2012), the skiers poled 1.25 times per breath at low efforts, and they also reported that BF becomes more synchronised with PF as the workload increases. Therefore, it appears that in the early intensities either a synchronised 1:1 ratio or a desynchronised 1:1.25 ratio is used to result in a 1:1 ratio at the higher intensities. In our study this was the case for most of the subjects (seven out of 10), who used a clear 1:1 BF to PF ratio from the very beginning or started with a BF<PF relationship to soon adopt a 1:1 ratio (3 subjects). However, this was not the case for the other three skiers, who adopted previously undocumented desynchronised BF-PF cross-like patterns. Therefore, our initial hypothesis stating that there would be a high correlation between BF and PF has been confirmed for most of the subjects when performing the V1 test as well, which would mean

that the assessment of HRVTs using the tendency of fHF would most likely be unsuccessful.

The DS test results can be compared to the studies of Holmberg and Calbet. (2007) in DS, Mourot et al. (2014) in ski-mountaineering, and Faria (1994) in simulated bilateral asynchronous poling. These studies showed a BF to PF ratio different than a 1:1 ratio with a tendency of a 1:1.1 ratio. Based on these results, and with the observation that the importance of the upper-body in the locomotion is smaller during DS, we hypothesised a weaker BF-PF entrainment for the DS test. Our results once more partly verified this as seven subjects adopted a desynchronised BF<PF relationship in the initial stages. However, four out of these seven subjects later adopted a previously undocumented BF>PF relationship (cross-like pattern) and the other three adopted a synchronised 1:1 ratio. Moreover, one other skier adopted a synchronised 1:1 BF to PF ratio throughout the test which then changed to a desynchronised BF>PF relationship in the last intensities. Another subject adopted an interesting BF>PF relationship throughout the entire test. Therefore, our initial hypothesis stating that there would be a high correlation between BF and PF, despite the results of Faria (1994), has only been partly confirmed for some of the subjects when performing the DS test. This means that the assessment of HRVTs using the tendency of fHF would probably be unsuccessful for some subjects.

To the best of our knowledge, this cross-like BF-PF pattern adopted by some subjects in the V1 and DS tests where BF becomes higher than PF in the highest workloads has not been reported in the literature. Taking into account that this phenomenon only occurs in the last few stages and that the last stage might be unfinished because the tests goes until failure, we propose that, due to the workload difficulties, certain subjects are un-able to keep increasing their tV, causing a steeper increase in BF. This would prevent the BF to become synchronised with PF, forcing the subjects to adopt a less convenient over proportional increase in BF in order to be able to keep up with the increasing workload. One might think that this would represent a problem when dividing the HFP_{LOC} and HFP_{RSA} components, as HFP_{LOC} would appear in a lower frequency than HFP_{RSA}, which would prevent in setting a cut-off frequency in PSD. Nevertheless, this is not the case in the present study, because it does not propose a problem in the assessment of $HRVT_{2\text{-}RSA}$, which is explained as follows. In the DS test the very high rate of movement, which is a consequence of the required bilateral asynchronous poling, leans towards the appearance of the HFP_{LOC} component at higher frequencies than the HFP_{RSA}. In regards to the DP test, $HRVT_{2\text{-}RSA}$ was not assessed all the same, due to the high entrainment between BF and PF, which has nothing to do with the cross-like pattern. In any case, as in the DP test the frequencies of both variables were very close, the two components were merged instead of the HFP_{LOC} component appearing at a lower frequency than the HFP_{RSA} component.

In regards to the NW test, the results were expected to be similar to the DS test, the reason for that is that NW mimics DS, and thus, the techniques are very similar. Our results showed that the subjects adopted two different BF-PF relationships. On one hand, four subjects showed no BF-PF synchronisation (BF<PF) but their frequencies got closer to each other with the increase in intensity, which is in line with what Faria (1994) reported. On the other hand, six subjects used synchronised or partly synchronised patterns, as they all started with a 1:2 BF to PF ratio and later changed to a 1:1 ratio or at least got closer to a 1:1 ratio. Therefore, it seems that the assessment of HRVTs using the tendency of fHF would most likely only be unsuccessful for some subjects.

Therefore, if we compare these results of the present study regarding the relationship between BF and PF with workload increase to our initial hypotheses, the agreement is in line for most of the subjects for the five tests. Being test-specific, an especially high BF to PF ratio was confirmed for the DP test, but also in the V1 test, as a 1:1 BF to PF ratio was predominant among the subjects. In the V2 test, a high BF-PF correlation was also confirmed as most of the subjects adopted 1:1 and/or 1:2 BF to PF ratios, whereas in the

DS and NW tests, the BF-PF entrainment was confirmed to be lower. This would entail, that, if the assessment of HRVTs using the tendency of fHF would be attempted, it would most likely be unsuccessful. The unsuccessfulness would be due to the assumption of linearity in the PF-workload trend, which would then entail a linearity in the BF-workload trends.

If we only focus on the relationship of PF with workload instead of the relationship between BF and PF with workload increase, we can see that our hypothesis was correct, as our results showed that the relationship between PF and workload was very linear for the V1 test and quite linear for the other tests (excluding the NW test). In the NW test, there was a change in the trend in the last few stages, as the constant increase of PF with workload suddenly suffers an inflection point. This could be explained by the observation that the change in the PF-workload trend corresponds to the phase when the subjects changed from walking to running.

Finally, focusing on the relationship between BF and workload, the results of the present study showed that, even if the exponential regression line always explains a greater percentage of the total variability than what the linear regression line does, the linear line also explains very high amounts of the total variability. This was especially true for the DP test, which agrees with our initial hypothesis, i.e., we hypothesised that a 1:1 BF to PF ratio along with the linearity of the PF-workload trend would bring a linear BF-workload trend. However, the results also showed that the least linear BF-workload trend among the tests was the V1 test. This was unexpected because in a similar fashion as the DP test, it was unexpected that the V1 test presents an especially high BF-PF coupling in direct consequence of a high dependence of the upper body in propulsion. Thus, the BF-workload trend of the V1 test was expected to be the second most linear instead of being the least linear, whereas the BF-workload trends of the NW and DS tests were expected to be the least linear because of the fact that they presented the lowest dependence on the upper-body for propulsion. It is difficult to speculate what could be the reason behind the poor BF-workload linearity, especially since an absence of PF-workload linearity would be the most natural explanation. However, this is not the case as the V1 test presents the most linear PF-workload trend among the five analysed tests. Nevertheless, at the same time, the almost perfect PF-workload linearity leads us, by process of elimination, to conclude that the relationship between BF and PF should be the cause. We could speculate that, apart from the BF-PF cross-like patterns adopted by three skiers, the BF<PF patterns adopted in the early stages by the other three subjects before synchronizing to a 1:1 BF to PF ratio are also to blame. This seems to be the only possible explanation strong enough to get the unexpected mean of the BF-workload trend, the explanation is as follows: only by summing up the subjects experiencing one of these two nonlinear or partly nonlinear BF-PF relationships (altogether six out of ten subjects) it is possible to get enough weight to unbalance the mean of the BF-workload trend towards a nonlinearity, to the detriment of the 1:1 BF to PF ratio linearity of the rest of the four subjects.

In summary, both the relative linearity between PF-workload and BF-workload trends has been supported by our results for the five different tests. Thus, the assessment of HRVTs using the tendency of fHF would most likely be unsuccessful in most cases and in any of the five tests, especially in the DP test.

6.3 Assessment of Ventilatory and Heart Rate Variability Thresholds

The determination of both VT_1 and VT_2 based on the behaviour of different ventilatory gas exchange variables with workload increase was quite challenging, as the plots did not always clearly show the inflection points in the same two stages. Nevertheless, all VTs, except VT_1 of one subject's DP test were managed to be determined. These difficulties were expected, as the only researcher involved in the assessment of VTs had limited experience in this field. Another fact was that the study of Fabre et al. (2012) already concluded that during the V2 test, the determination of VT_2 from ventilatory gas analysis must be used cautiously because the PF entrains BF, preventing the usual inflection-like behaviour of the ventilatory variables. Based on this, it was hypothesised that the assessment of VTs from ventilatory gases would be challenging, and therefore, it might not be that successful in terms of validity.

Regarding the assessment of HRVTs, our study failed to assess both thresholds from any of the four methodologies in five maximal tests because their data were discarded due to excess of artefacts as previously mentioned. This could have probably been solved by using an ECG instead of a HR monitor, as practical problems like the improper placement of the HR belt were discarded based on the experience of both the researchers and the subjects. Moreover, $HRVT1_{1\text{-}MSD}$ could not be assessed in one of the DS tests because of the abnormal behaviour of its trend. Regarding $HRVT_{2\text{-}HFP\text{-}RSA}$, it was not assessed in DP and V1, because as it was mentioned in the methods section: The strong synchronisation between PF and BF provoked a single merged peak in the HF band that prevented a differentiation of the HFP_{RSA} component from it. This was not unexpected, because based on literature we previously hypothesised that dividing these two spectral components would probably not be possible in these two tests, especially in the DP test. However, besides these difficulties, the rest of HRVTs were assessed, including the VT_1 that could not be assessed from the ventilatory method, meaning that, in this particular case, the HRV based methods were preferable to the conventional ventilatory method. Nevertheless, this does not mean that the assessment of the submaximal thresholds from the tested HRV-derived methods is more preferable to their assessment from the more conventional ventilatory methods, as, these HRV based methods are more time demanding and require more knowledge, and the validity and reliability on XC skiers has not been discussed yet.

When assessing the $HRVT_{2\text{-}HFP\text{-}RSA}$, it was noted that in every modality there was a certain tendency for the synchronisation of the two spectral components (HFP_{RSA} and HFP_{LOC}) in the high intensities. This is in line with our hypothesis, based on the study of Fabre et al. (2007) in V2 and V2-alternate skating. This study reported an increased BF-PF synchronisation with workload increase, from which it was deduced that the same synchronisation phenomenon would be reflected in their corresponding spectral components.

Being test-specific, in V2 two clear peaks emerged as it was reported in the literature (Blain et al. 2009, Di Michele et al. 2012 & Mourot et al. 2014). However, in the DS tests, several peaks emerged, and even if the main peak corresponded to RSA, the peak corresponding to LOC was the last one appearing, which was often non prominent. To the best of our knowledge, this has been undocumented to this date, and the explanation of both the origin of the minor peaks unrelated to RSA or LOC, and the non prominence of the LOC peak, is beyond our knowledge. Moreover, it is important to mention that, because pfHF was often a non prominent peak, it would not be possible to recognize it unless the PF was timed, which would mean that it would not be possible to set a cut-off frequency between HFP_{RSA} and HFP_{LOC}. Therefore, the assessment of $HRVT_{2\text{-}HFP\text{-}RSA}$ in the DS test requires the timing of PF.

Nevertheless, even if in these two tests there was always a peak truly representing the LOC, this was not the case in the NW test. This was surprising considering that NW being an exercise mode created to mimic DS would entail relatively similar locomotions between NW and DS. Even if in these spectras, as expected from the literature, two main peaks emerged and the first one corresponded to RSA, instead of the later prominent peak corresponding to the double of PF, its frequency was similar to PF, which would correspond to the half of the rate of arm movements. This would mean that LOC is not based on each poling action, but on the whole poling cycle (either on the left or on the right poling action), and not to each poling action as reported in Di Michele et al. (2012) and also in the rest of our results, in what could be considered a new LOC mode. We proposed that the appearance of this novel LOC mode and the failure to find a peak at the double at PF must have been related to the distinctive nature that NW was the only test that did not use roller skies. We proposed that, the absence of roller skies prevents the existence of a rolling/sliding phase, consequently reducing the stride length, which demands a higher PF for maintaining a given workload. This would be responsible for increasing the rate of arm movements higher than the upper frequency limit of PSD (2 Hz), which would explain the failure to find a peak at the rate of arm movements. Also, the absence of roller skies, apart from a higher PF, most likely results in lower arm forces, which might also have something to do with the appearance of the new cardilocomotor coupling mode.

Regarding the appearance of non prominent peaks that seem to be unrelated to RSA and LOC, we proposed that it could be related to the use of a bilateral asynchronous poling, because the appearance of these peaks only occurred in DS and NW, the only techniques using this kind of poling action. This explanation seems reasonable, especially as the bilateral desynchronisation entails a more constant locomotion with more movements of a less explosive nature, which would be more homogeneously spread in PSD instead of being focused in just a few frequencies.

In the three aforementioned modalities, a common phenomenon occurred: there was a tendency for the synchronisation of PF with BF an increase of exercise intensities, already reported by Fabre et al. (2007a). However, the remaining two modalities (V1 & DP) are the only modalities in this study that present a unique upper-body/lower-body locomotion (full body moves at the same frequency), which makes the RSA-LOC coupling stronger. Consequently, there is a strong tendency for a single spectral peak from the earliest stages in the HF band that makes it impossible to obtain HFP_{RSA}. Nevertheless, this was not unexpected as it was previously proposed as a hypothesis.

In regards to the possible origin of the LOC, it has been proposed to be a modulation of venous return by limb muscle contraction in cycling (Blain et al. 2009), a reasoning that also seems plausible for the five XC-skiing related techniques of the present study. Nevertheless, what is sure is that the LOC needs to have an either neural, mechanical or neuromechanical origin.

In summary, the HFP_{RSA} and HFP_{LOC} patterns of the studied tests excluding the NW test, were not unexpected as they closely corresponded to the frequencies of BF and PF. On one hand, in every test beside the NW test, the appearance of HFP_{LOC} was synchronised to the rate of the arm movements, which was the same as PF in V2, V1, and DP, but was the double of PF in DS. On the other hand, in the NW test, the frequency of the HFP_{LOC} was closely related to PF, which corresponds to the half of the rate of arm movements, and thus, to the whole poling cycle instead of each poling action. Thus, during these four techniques using roller skies, pfHF closely corresponded to the rate of arm movements (i.e $2 \cdot PF$ in DS and PF in V2, V1 and DP), independently of the frequency of the first peak (i.e. fHF), whereas in NW it corresponded to half of it. Therefore, the present results support the evidence that the breathing pattern and the poling action modulate HR during the analysed XC skiing techniques, and reports a new form of LOC.

6.4 Agreement between Ventilatory and Heart Rate Variability Thresholds

Regarding the agreement between VT_1 and the two different equivalents assessed from time-domain analysis of HRV, the agreements were not as good as expected. The only test where $HRVT_{1-STD}$ was a proper assessor of VT_1 was the DS test, as the $HRVT_{1-STD}$ values obtained in the rest of the tests (i.e. DP, NW, V1 and V2) failed to successfully assess VT_1. The mean VT_1 and $HRVT_{1-STD}$ of the DS test were strongly correlated ($r = 0.772$), and their means were not significantly different with virtually no bias. However they do present a relatively wide agreement interval and $HRVT_{1-STD}$ explained 59.5 % of the total variability of VT_1. If we compare this to the results obtained in similar studies, the cycle-ergometer based study of Karapetian et al. (2008) and the walking study of Dourado et al. (2010) give similar results, showing even better agreements. In the first study the mean VO_2 values at $HRVT_{1\text{-}STD}$ and VT_1 were not significantly different either, they had no bias and they were very highly correlated (r=0.98). The later study also showed insignificant mean differences between the VO_2 values at the two thresholds, which were also very strongly correlated (r=0.946). Moreover, they also reported that the Bland-Altman plot revealed agreement, that the confidence interval was wide and that the linear regression explained up to 89.6%.

In regards to the other HRV based method for the assessment of VT_1, which was named $HRVT_{1\text{-}MSD}$, was a relatively good assessor in the V2 and DS tests. In the V2 and DS tests, these two variables were strongly correlated (r = 0.765 and 0.752, respectively), and their means were not significantly different, even though their negative biases (-3 and -6 bpm, respectively) did show that there was a slight underestimation of $HRVT_{1\text{-}MSD}$ over VT_1. The regression line explained 58.51 % and 56.49 % of the total variability, respectively, and the limits of agreement were not especially narrow in either of the tests. If we compare these results with the results obtained in the study of Karapetian et al. (2008), which is, to the best of our knowledge, the only other study using $HRVT_{1\text{-}MSD}$, it can be concluded that the results of the present study did not show as good agreement. In the referred study, apart from the fact that the mean VO_2 values at $HRVT_{1\text{-}MSD}$ and VT_1 were not significantly different, they had no bias and they were very strongly correlated (r=0.98).

The $HRVT_{1\text{-}MSD}$ values of the present study that were obtained in the rest of the tests (i.e. DP, NW and V1) failed to successfully assess VT_1. Moreover, it is also remarkable that, in the study of Karapetian et al. (2008) where both $HRVT_{1\text{-}STD}$ and $HRVT_{1\text{-}MSD}$ were assessed, the two thresholds corresponded to the same workload in every case, whereas this was not true in the present study.

In regards to VT_2, it seems that $HRVT_{2\text{-}HFP}$ was not a proper assessor in either of the tests, on the contrary to the close agreements reported for swimming (Di Michele et al. 2012) and ski-mountaineering (Mourot et al. 2014). However, it seems that $HRVT_{2\text{-}HFP\text{-}RSA}$ was a good assessor of VT_2 in the NW test and an okay assessor in the DS test. Other than that, the $HRVT_{2\text{-}HFP\text{-}RSA}$ values obtained in the rest of the tests (i.e. DP, V1 and V2) failed to successfully assess VT_2. In the NW and DS tests, the two variables were strongly correlated (r = 0.818 and r = 0.664, respectively) and their means were not significantly different. However, if in the NW test there was virtually no bias, in the DS test, the -9 bpm bias represented a considerable underestimation of $HRVT_{2\text{-}HFP\text{-}RSA}$ over VT_2. Moreover, the DS test presented a wider agreement interval than that of the NW test (-35 to 17 bpm and -14.9 to 13.8 bpm, respectively) and the regression line explained a smaller amount of the total variability (44.1 and 48.3 %, respectively).

Comparing our results with the results obtained in the literature, the swimming study of Di Michele et al. (2012) reported a high correlation between LT_2 and $HRVT_{2\text{-}HFP\text{-}RSA}$ (r=0.93) where the insignificantly different means showed a low bias (-1 bpm) and a very narrow agreement interval (-7 to 5 bpm). The ski-mountaineering study of Mourot et al. (2014) reported a -3 bpm bias corresponding to insignificantly different means that were largely correlated (r=0.647), with limits of agreement -22 to 14 bpm. Therefore, the results obtained in our NW test were similar to the results obtained in these two studies, because all of them have small biases and moderately narrow limits of agreement. On the contrary, the agreement between $HRVT_{2\text{-}HFP\text{-}RSA}$ and VT_2 was not as good in the DS test, and probably, the use of $HRVT_{2\text{-}HFP\text{-}RSA}$ as assessor of VT_2 would not be valid enough for applicability in the field because of the wide agreement interval due to a wide inter-individual difference as well as a -9 bpm bias which often exceeds the daily variations in HR. The study of Lamberts and Lambert (2009) supports this statement, as it was concluded that a change in HR recovery of more than 6 bpm or a change in submaximal HR of more than 3 bpm can be regarded as a meaningful change under controlled conditions. Finally, it is also important to note that in our results, in the same way that it happened in the aforementioned literature, when $HRVT_{2\text{-}HFP\text{-}RSA}$ was determined, it was a better assessor of VT_2 than $HRVT_{2\text{-}HFP}$.

6.5 Limitations and Practical Applications

Despite our efforts in an attempt to achieve excellence in the present study, it is not exempt of limitations. First of all, there is the issue of the limited amount of subjects, which is limiting in obtaining strong enough statistical conclusions. However, recruiting more subjects that fit the right profile, apart from being time and effort demanding, it would also increase the financial budget of the study. Moreover, the selected method for recording the RR intervals reduced the sample size to a further extent in some cases. This problem could be relatively easily solved if the RR intervals would have been recorded from ECG instead of the HR monitors. In the present study, the RR extraction was performed directly and some recordings could not be used (five out of 50) because of an excessive number of artefacts. Thus, further studies should be conducted with ECG or ECG-like instruments.

Another limitation is that the assessment of $HRVT_2$ requires a particularly high level of understanding and expertise in study fields that are quite distant to exercise physiology, such as spectral analysis and complex mathematical modelling. There might also be room for improvement in the design of the incremental test protocols. The decision of using protocols with short brakes was taken in order to obtain lactate samples (data that was not used in the present study). However, the use of a non-stop protocol would have been preferable, because it would enable us to use a time varying spectral analysis method instead of a non-time varying analysis. Even if this would require more work, it would allow more accurate determination of HRVTs. Moreover, it would have also been preferable to design the protocols to have more stages (approx. eight to ten) by either making them shorter or, preferably, making the increases milder, in speed and/or percent inclination.

The wisdom of the decision of determining VTs at stage values instead of determining in a theoretical midpoint between two stages could also be debated. This decision was taken in order to match the VTs and HRVTs easier, as the HRVTs could only be determined at stage values. Nevertheless, when talking about the HR values, the possible error made when selecting a stage value as VT value instead of a theoretical inter-stage value where VT would lie, probably does not make the data unavailable for the daily training. This deduction is based on the idea that, if the assessment were correct, the maximal difference between the selected stage value as VT value (closest stage to VT value) and the real VT value lying between two stages is of half a stage, representing a difference no bigger than ± 5 bpm, which would be often covered by the daily variations in HR. The results from the study of Lamberts and Lambert (2009) support this statement, as a range from 17 ± 8 to 6 ± 2 bpm and 8 ± 3 to 3 ± 1 bpm was found as a day-to-day change in HR recovery and submaximal HR, respectively.

The method used for the assessment of VTs may not have been the most optimal method. The average values of the last 90 seconds of different ventilatory data were used as stage values to match the timeframe used for the assessment of HRVTs, which, following the literature, was based on 90 second averages. However, we speculate that 60 second averages may have given better estimates of the stage value because it allows a longer (120 second) period for reaching the steady-state, which could result in a more precise assessment of VTs. The possibility for using continuous ventilatory data instead of stage value averages was discarded because the interrupted nature of the test protocols causes non-desired drops in the ventilatory parameters that complicates the visualization of the inflection points. However, using continuous data as breath-to-breath data or as average values of every 3-5 seconds after discarding the data from the first 90-120 seconds of each stage could have been a valid method.

From a practical point of view, the present study allows us to establish that, in the upper body based exercise disciplines like XC skiing, where a really tight coordination exists between BF, PF and workload, the determination of VTs and especially HRVTs, is a hazardous task that must be used cautiously. This is especially seen in the DP test, a technique where the entrainment between the locomotor and the breathing rhythms is the highest, whereas it is seen to be the least meaningful in the DS test, where the assessment of $HRVT_1$ from the two different time-domain methods was most successful. The assessment of $HRVT_{1\text{-}MSD}$ was also successful in the V2 test, and the assessment of $HRVT_2$ was only successful in the NW test. Thus, if the assessment of VTs from HRV is geared towards XC skiers, we propose to single out specific XC skiing protocols preferably using the DS technique to determine $HRVT_1$ and use the V2 test for the assessment of $HRVT_2$, rather than using all five techniques. As a matter of fact, Fabre et al. (2012) already suggested to use the DS test instead of the skating techniques for the determination of the VT_2. Nevertheless, the use of HRVTs in the DS and V2 tests requires further research and its applicability must be evaluated in training routines to set parameters for training intensity and monitoring training adaptations. Furthermore, the assessment of $HRVT_2$, apart from being time-demanding, requires high levels of knowledge, and thus, unless the methodology is simplified in the future, its applicability will be questioned. Therefore, as the assessment of VTs from HRV has been proven to be unsuccessful for most cases, alternative non-invasive and economical methods, such as heart rate deflection (Fabre et al. 2008) could also be proposed, but the use of such methods needs to be more profoundly evaluated as well.

6.6 Conclusion

BF and PF measured by HRV and their respective equivalents determined conventionally did not present any high variances, as it was expected, and if any, the HRV measurements mostly tended to slightly overestimate the reference measurements. Regarding the relationship between BF and PF with workload increase, three different kinds of BF-PF relationships were observed throughout the five tests regarding the mean values: 1) In the DP test, these two parameters were synchronised (1:1 BF to PF ratio) almost from the very beginning; 2) the V1 and DS tests showed statistical differences in early and late stages as the BF trend surpassed the PF trend in a cross-like pattern; and 3) the NW and V2 tests presented clearly different BF and PF values except at the two highest intensities. All these results agree with the literature except for the significantly different mean values that were found in the final stages of the V1 and DS tests that leads to the cross-like pattern, a phenomenon that, to the best of our knowledge, was unreported to the date. All the tests had the following phenomenons in common: a tendency for an increased BF-PF coupling with workload increase which was reported in Crooks et al. (2012), a higher PF than BF during the initial workloads, and the existence of significant inter-individual differences in the adopted BF-PF pattern. When the BF-PF patterns were analysed individually, we found out that our initial hypotheses were accurate. Being test-specific, an especially high BF to PF ratio was confirmed for the DP test, as well as also in the V1 test, as in both tests a 1:1 BF to PF ratio was predominant among the subjects. In the V2 test, a high BF-PF correlation was also confirmed as most of the subjects adopted 1:1 and/or 1:2 BF to PF ratios, whereas in the DS and NW tests, the BF-PF entrainment was confirmed to be lower. This would entail, that, if the assessment of HRVTs using the tendency of fHF would be attempted, it would probably be unsuccessful, because, assuming linearity in the PF-workload trend, it would entail linearity in the BF-workload trends.

Focusing on the relationship of PF with workload, our hypothesis was confirmed, as our results showed that the relationship between PF and workload was very linear for the V1 test and quite linear for the rest of the tests except for the NW test. The nonlinearity of the NW test seemed to be due to the switch from walking to running. When looking at the relationship between BF and workload, the results of the present study showed the linear regression line fit almost as good as the exponential regression line. Thus, the hypothesised linearity of the BF-workload trend was partially confirmed, and was best seen in the DP test exactly as hypothesised. However, the results also showed that the least linear BF-workload trend among the tests was the V1 test, which was unexpected. In summary, as the high entrainment between BF and PF and the relative linearity of both PF-workload and BF-workload trends have been supported by our results for the five different tests, we conclude that, the assessment of HRVTs using the tendency of fHF would most likely be unsuccessful in most of the subjects and any of the five tests. This is very likely to be seen in the DP test.

In regards to the determination of both VT_1 and VT_2 based on the behaviour of the ventilatory gases, presented a challenge as the plots did not always clearly show the expected two inflection points. Nevertheless, all VTs except VT_1 of one subject's DP test were managed to be determined. The numerous ways of entrainment of breathing to locomotion in addition to the physiological demands may be the most important reasons for these unexpected difficulties. Moreover, the only researcher involved in the assessment of VTs had limited experience in the field. The study of Fabre et al. (2012) concluded that during the V2 test, the determination of VT_2 from ventilatory gas analysis must be used cautiously. This observation was based on the fact that PF entrains BF, preventing the usual inflection-like behaviour of the ventilatory variables.

In regards to the assessment of HRVTs, the main research question of the present study, our study failed to assess both thresholds from any of the four methodologies in five maximal tests because the corresponding data were discarded due to excess of artefacts. Moreover, $HRVT1_{1\text{-}MSD}$ could not be assessed in one of the DS tests because of the abnormal behaviour of its trend. Regarding $HRVT_{2\text{-}HFP\text{-}RSA}$, it was not assessed in DP and V1, because the strong synchronisation between PF and BF provoked a single merged peak in the HF band that prevented to differentiate the HFP_{RSA} component from it. This was not unexpected as, based on the literature, we hypothesised that dividing these two spectral components would probably not be possible in these two tests, especially in the DP test. However, besides these difficulties, the rest of HRVTs were assessed, including VT_1 which could not be assessed from the ventilatory method, meaning that, in this particular case, the HRV based methods were preferable to the conventional ventilatory method. Nevertheless, this does not mean that the assessment of the submaximal thresholds from the tested HRV-derived methods is preferable to their assessment from the more conventional ventilatory methods, as these HRV based methods are more time demanding and require more knowledge. On top of that, the validity and reliability of the assessment of VTs from HRV in XC skiers has not been discussed yet.

If we go more in depth into the utilized spectral analyses for the assessment of $HRVT_{2\text{-}HFP\text{-}RSA}$, the HFP_{RSA} and HFP_{LOC} patterns of the studied test disregarding the NW test, were not unexpected as they closely corresponded to the frequencies of BF and PF. In every test beside the NW test, the appearance of HFP_{LOC} was synchronised to the rate of arm movements, which was the same as PF in V2, V1 and DP, but was $2 \cdot PF$ in DS because of the bilateral asynchronous nature of its poling action. However, in the NW test, the frequency of HFP_{LOC} was closely related to PF, which corresponds to the half of the rate of arm movements, and thus, corresponds to the whole poling cycle instead of each poling action. Thus, during these four techniques using roller skies, the pfHF closely corresponded to the rate of arm movements (i.e $2 \cdot PF$ in DS and PF in V2, V1 and DP), independently of the frequency of the first peak (i.e. fHF), whereas in NW pfHF corresponded to half of the rate of arm movements. Therefore, these results support the evidence that the breathing pattern and the poling action modulate HR during the analysed XC skiing techniques, and the also report a previously undocumented form of LOC for the NW test.

The agreement between VT_1 and the two different equivalents assessed from time-domain analysis of HRV was done by comparing HR values. These agreements were not as good as expected. The only test where $HRVT_{1\text{-}STD}$ showed a proper assessment of VT_1 was during the DS test, and $HRVT_{1\text{-}MSD}$ was a relatively good assessor of VT_1 only in the V2 and DS tests. In the rest of the tests, both $HRVT_{1\text{-}STD}$ and $HRVT_{1\text{-}MSD}$ failed to successfully assess VT_1. In the NW test and especially in the V1 test, $HRVT_{1\text{-}MSD}$ were very different to VT_1. Moreover, it is also remarkable that, in the cycle ergometer based study of Karapetian et al. (2008) where both $HRVT_{1\text{-}STD}$ and $HRVT_{1\text{-}MSD}$ were assessed, the two thresholds corresponded to the same workload in every case, whereas this was not the case in the present study. Regarding VT_2, where the agreement was also tested based on the HR values, it seems that $HRVT_{2\text{-}HFP}$ was not a proper assessor of VT_2 in either of the tests, contrary to the fact that there were close agreements reported in swimming (Di Michele et al. 2012) and ski-mountaineering (Mourot et al. 2014). However, it seems like $HRVT_{2\text{-}HFP\text{-}RSA}$ was at least a good assessor of VT_2 in the NW test and an okay assessor in the DS test, even though it was not a valid assessor in the rest of the tests. Nevertheless, it was concluded that $HRVT_{2\text{-}HFP\text{-}RSA}$ would probably not be a valid enough assessor of VT_2 regarding the applicability in the field, because, apart from a wide agreement interval, its underestimation exceeds the magnitude of the daily variations in HR. Finally, it is also important to note that in our results, similar to the aforementioned literature, when $HRVT_{2\text{-}HFP\text{-}RSA}$ was determined, it was a better assessor of VT_2 than $HRVT_{2\text{-}HFP}$.

In conclusion, the present study shows results supporting that the assessment of VTs, and in particular, the assessment of HRVTs in the upper body based locomotions seen in XC skiing, is a task that requires expertise. Moreover, the assessment of HRVTs also requires the selection of an appropriate method, which, might not always ensure the successful determination of HRVTs. However, one would encourage the use of both time-domain HRV methods for the assessment of VT_1 in the DS test, and the use of $HRVT_{1\text{-}MSD}$ in the V2 test, because of their simplicity. The use of the more complex frequency-domain HRV methods are only encouraged for experimental settings, especially as further research is needed to prove their validity and reliability. Nevertheless, it seems that $HRVT_{2\text{-}HFP\text{-}RSA}$ is a valid assessor of VT_2 in the NW test.

7 REFERENCES

Achten, J. & Jeukendrup, A. E. 2003. Heart Rate Monitoring. Applications and Limitations. Sports Medicine. 33 (7), 517-538.

Anosov, O., Patzak, A., Kononovich, Y. & Persson, P. B. 2000. High-frequency oscillations of the heart rate during ramp load reflect the human anaerobic threshold. Eur J Appl Physiol. 83, 388-394.

Aubert, A. E., Seps, B. & Beckers, F. 2003. Heart Rate Variability in Athletes. Sports Medicine. 33 (12), 889-919.

Blain, G., Meste, O. & Bermon, S. 2005a. Influences of breathing patterns on respiratory sinus arrhythmia in humans during exercise. Am J Physiol Heart Circ Physiol. 288, 337-895.

Blain, G., Meste, O., Blain, A. & Bermon, S. 2009. Time-frequency analysis of heart rate variability reveals cardiolocomotor coupling during dynamic cycling exercise in humans. American Journal of Physiology - Heart and Circulatory Physiology. 296, 1651-1659.

Blain, G., Meste, O., Bouchard T. & Bermon, S. 2005b. Assessment of Ventilatory Thresholds during graded and maximal exercise test using time-varying analysis of respiratory sinus arrhythmia. Br J Sport Med. 39, 448-452.

Bosquet, L., Papelier, Y., Léger, L. & Legros, P. 2003. Night heart rate variability during overtraining in male endurance athletes. Journal of Sports Medicine and Physical Fitness. 43, 506-512.

Buchheit, M., Simon, C., Piquard, F., Ehrhart, J. & Brandenberger, G. 2004. Effects of increased training load on vagal-related indexes of heart rate variability: a novel approach. Am J Physiol Heart Circ Physiol. 287, 2813-2818.

Buchheit, M., Solano, R. & Millet, G. M. 2007. Heart-Rate Deflection Point and the Second Heart-Rate Variability Threshold During Running Exercise in Trained Boys. Pediatric Exercise Science. 19, 192-204.

Cottin, F., Lepretre, P.-M., Lopes, P., Papelier, Y., Médigue, C., & Billat, V. 2006. Assessment of Ventilatory Thresholds from Heart Rate Variability in Well-Trained Subjects during Cycling. International Journal of Sports Medicine. 27, 959-967.

Cottin, F., Médigue, C., Lepetre, P.-M., Papelier, Y., Koralsztein, J.-P. & Billat, V. 2004. Heart Rate Variability during Exercise Performed below and above Ventilatory Thresholds. Medicine & Science in Sports & Exercise. 4, 594-600.

Cottin, F., Médigue, C., Lopes, P., Lepretre, P.-M., Heubert, R. & Billat, V. 2007. Ventilatory Thresholds Assessment from Heart Rate Variability during an Incremental Exhaustive Running Test. International Journal of Sports Medicine. 28, 287-294.

Crooks, S., Killick, A., Vetsch, S. & Herzog, W. 2012. Rate and Time of Breath in One and Two Skate Cross Country Skiing. Journal of Undergraduate Research in Alberta. 2 (1), 6.

Di Michele, R., Gatta, G., Di Leo, A., Cortesi, M., Andina, F., Tam, E., Da Boit, M. & Merni, F. 2012. Estimation of the Anaerobic Threshold from Heart Rate Variability in an Incremental Swimming Test. Journal of Strength and Conditioning Association. 26 (11), 3059-3066.

Dourado, V. Z., Banov, M. C., Marino, M. C., de Souza, V. L., de O. Antunes, L. C. & McBurnie, M.A. 2010. A Simple Approach to Assess VT During a Field Walk Test. Int J Sports Med. 31, 698-703.

Durnin, J. V. G. A. & Rahaman, M. M. 1967. The assessment of the amount of fat in the human body from measurements of skinfold thickness. Br J Nutr. 21, 544-550.

Fabre, N., Bortolan, L., Pellegrini, B., Zerbini, L. Mourot, L. & Schena, F. 2012. Anaerobic Threshold Assessment Through the Ventilatory Method During Roller-Ski Skating

Testing: Right or Wrong? Journal of Strength and Conditioning Research. 26 (2), 381-387.

Fabre, N., Passelergue, P., Bouvard, M. & Perrey, S. 2008. Comparison of Heart Rate Deflection and Ventilatory Threshold During a Field Cross-Country Roller Skiing Test. Journal of Strength and Conditioning Research. 22 (6), 1977-1984.

Fabre, N., Perrey, S., Arbez, L. & Rouillon, J.-D. 2007a. Paced Breathing in Roller-Ski Skating: Effects of Metabolic Rate and Poling Forces. International Journal of Sports Physiology and Performance. 2, 46-57.

Fabre, N., Perrey, S., Arbez, L. & Rouillon, J.-D. 2007b. Neuro-mechanical and chemical influences on locomotor respiratory coupling in humans. Respiratory Physiology and Neurobiology. 155, 128-136.

Faria, I. E. 1994. Ventilatory response pattern of Nordic skiers during simulated poling. J Sports Sci. 12:255-259.

Freeman, J. V., Dewey, F. E., Hadley, D. M., Myers, J. & Froelicher, V. F. 2006. Autonomic Nervous System Interaction With the Cardiovascular System During Exercise. Progress in Cardiovascular Diseases. 48 (5), 342-362.

Hainsworth, R. 1998. Chapter one: Physiology of the Cardiac Autonomic System. In: Clinical Guide to Cardiac Autonomic Tests edited by Malik, M. Kluwer Academic Publishers. 3-28.

Hartikainen, J. E. K., Tahvanainen, K. U. O. & Kuusela, T. A. 1998. Chapter six: Short-Term Measurements of Heart Rate Variability. In: Clinical Guide to Cardiac Autonomic Tests edited by Malik, M. Kluwer Academic Publishers. 149-176.

Holmberg, H. C., Lindinger, S., Stöggl, T., Eitzlmair, E. & Müller, E. 2005. Biomechanical Analysis of Double Poling in Elite Cross-Country Skiers. Medicine & Science in Sports & Exercise. 807-818.

Holmberg, H. C. & Calbet, J. A. L. 2007. Insufficient ventilation as a cause of impaired pulmonary gas exchange during submaximal exercise. Respiratory Physiology and Neurobiology. 157, 348-459.

Hopkins, W. G., Marshall, S. W., Batterham, A. M. & Hanin, J. 2009. Progressive statistics for studies in sports medicine and exercise science. Med Sci Sports Exerc. 41 (1). 3-13.

Hynynen, E., Uusitalo, A., Konttinen, N. & Rusko, H. 2006. Heart Rate Variability during Night Sleep and after Awakening in Overtrained Athletes. Medicine and Science in Sports & Exercise. 38, 313-317.

James, N. W., Adams, G. M. & Wilson, A. F. 1989. Determination of anaerobic threshold by ventilatory frequency. Int J Sports Med. 10, 192-196.

Jones, A. M. & Doust, J. H. 1998. Assessment of the lactate and Ventilatory Thresholds by breathing frequency in runners. Journal of Sports Sciences. 16, 667-675.

Kaikkonen, P., Hynynen, E., Mann, T., Rusko, H. & Nummela, A. 2012. Heart rate variability is related to training load variables in interval running exercises. Eur J Appl Physiol. 112, 829-838.

Karapetian, G. K., Engels, H. J. & Gretebeck, R. J. 2008. Use of Heart Rate Variability to Estimate LT and VT. Int J Sports Med. 29, 652-657.

Kindermann, W., Simon, G. & Keul, J. 1979. The significance of the aerobic-anaerobic transition for the determination of work load intensities during endurance training. Eur J Appl Physiol. 42, 25-34.

Lamberts, R. P. & Lambert, M. I. 2009. Day-to-Day Variation in Heart Rate at Different Levels of Submaximal Exertion: Implications for Monitoring Training. Journal of Strength and Conditioning Research. 23 (3), 1005-1010.

Meyer, T., Lucía, A., Earnest, C. P. & Kindermann, W. 2005. A conceptual Framework for Performance Diagnosis and Training Prescription from Submaximal Gas Exchange Parameters - Theory and Application. Int J Sports Med. 26, 38-48.

Mourot, L., Fabre, N., Savoldelli, A. & Schena, F. 2014. Second Ventilatory Threshold From Heart-Rate Variability: Valid When the Upper Body Is Involved? Int J of Sports Physiology and Performance. 9, 695-701.

Saalasti, S., Seppänen, M. & Kuusela, A. 2004. Artefact Correction for Heartbeat Interval Data. Advanced Methods for Processing Bioelectrical Signals. Available: Http://www.firstbeat.com/userData/firstbeat/download/saalasti_et_al_probisi_2004_congress.pdf

Task Force 1996. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. European Heart Journal. 17, 354-381.

Taylor, J. A., Carr, D. L., Myers, C. W. & Eckberg, D. L. 1998. Mechanisms Underlying Very-Low-Frequency RR-Interval Oscillations in Humans. Circulation, Journal of the American Heart Association. 98, 547-555.

Tulppo, M. 1998. Heart Rate Dynamics during Physical Exercise and during Pharmacological Modulation of Autonomic Tone. Oulu University Press.

Tulppo, M. P., Makikallio, T. H., Seppanen, T., Laukkanen, R. T. & Huikuri, H. V. 1998. Vagal modulation of heart rate during exercise: effects of age and physical fitness . Am J Physiol. 274, 424-429.

Tulppo, M. P., Makikallio, T. H., Takala, T. E. S., Seppänen, T. & Huikuri, H. V. 1996. Quantitative beat-to-beat analysis of heart rate dynamics during exercise. Am J Physiol. 271, 244-252.

Wasserman, K. & McIlroy, M. B. 1964. Detecting the threshold of anaerobic metabolism in cardiac patients during exercise. Am J Cardiol. 14, 844-852.

Wasserman, K., Whipp, B. J., Koyl, S. N. & Beaver, W. L. 1973. Anaerobic threshold and respiratory gas exchange during exercise. J Appl Physiol. 35, 236-243.

Westhoff, M., Rühle, K. H., Greiwing, A., Schomaker, R., Eschenbacher, H., Siepmann, M. & Lehnigk, B. 2013. Positional paper of the German working group "cardiopulmonary exercise testing" to ventilatory and metabolic (lactate) thresholds. Deutsche Medizinische Wochenschrift. 138, 275-280.

Yamamoto, Y., Hughson, R. L. & Nakamura, Y. 1992. Autonomic nervous system responses to exercise in relation to ventilatory threshold. Chest. 101, 206–210.