DETERMINATION OF HEART RATE VARIABILITY THRESHOLDS AND THEIR ASSOCIATIONS WITH LACTATE THRESHOLD IN NOVICE FEMALE RUNNERS

Iida Mustakoski

Master's Thesis in Exercise Physiology
Faculty of Sport and Health Sciences
University of Jyväskylä
Autumn 2023

ABSTRACT

Mustakoski, I. 2023. Determination of Heart Rate Variability Thresholds and Their Association with Lactate Threshold in Novice Female Runners, Faculty of Sport and Health Sciences, University of Jyväskylä, Master's thesis, 70 pp. 3 appendices.

Endurance type exercise at different intensities induces various physiological changes in human body. For desirable training stimuli the exercise intensity should be chosen according to what are the physiological changes at which training is currently targeted. Exercise thresholds (ET) are intensity bounds that are separating the exercise intensities with different physiological functions. ETs can be determined from measurements conducted during graded exercise test. Traditionally, blood lactate concentration and/or ventilatory functions are measured and thresholds are determined based on changes seen on those parameters. For lactate and ventilatory measurements, applicable equipment and experienced testing personnel are needed. Thus, new methods which are more easily executed and noninvasive would be useful. The purpose of this study was to test the applicability of HRV threshold methods presented in previous studies in nonathletic, healthy females.

In recent years, also other less invasive methods and physiological measures for threshold determination have been developed and tested. One of those is heart rate variability (HRV). In this study, the identifiability and reliability of heart rate variability thresholds (HRVT) compared to lactate threshold (LT) were tested. Lactate thresholds were determined according to the Finnish recommendations. Four methods were used for first heart rate variability threshold (HRVT1) and three methods for second heart rate variability threshold (HRVT2) determination. For HRVT1 determination RMSSD parameter of time domain analysis, HF power of frequency domain analysis, SD1 of Poincaré plot and DFA-a1 of detrended fluctuation analysis were used. Respectively, for HRVT2 determination HF power, SD2 and DFA-a1 were examined. Heart rate and speed at HRVT1 and HRVT2 determined by different methods were compared to HR and speed at LT1 and LT2, correspondingly. All the participants were adult female novice runners. There is currently a couple of previous studies investigating the reliability of heart rate variability thresholds. However, very few of them are conducted in females thus the results of the present study widen the knowledge about the applicability of HRVT methods among female population.

According to the results, HRVTs are mostly identifiable in novice female runners. The only exception was HRVT2 determined using HF power which could only be identified in half of the participants. Considering the analysis methods used (correlations, paired differences, ICCs, linear regressions, Bland Altman -plots) the most reliable HRVT1 method for LT1 estimation seemed to be the one based on time domain analysis and its RMSSD variable. None of the HRVT2 methods used appeared superior to others and all showed quite low correlations and reliability when compared to LT2. In the future, it would be important to conduct more studies of this topic in female population with different and maybe more strictly standardized testing protocols. The applicability of HRVTs in practical training would also be relevant topic to be investigated.

Key words: heart rate variability, heart rate variability threshold, lactate threshold, exercise test, endurance performance

ABBREVIATIONS

ANS	autonomic nervous system
BF	breathing frequency
CI	confidence interval
DFA-a1	detrended fluctuation analysis alpha 1
ECG	electrocardiogram
ET	exercise threshold
GXT	graded (maximal) exercise test
HF	high frequency (component of HRV power spectrum; 0,15–0,5 Hz)
HF*fHF	mathematical product of absolute high frequency power (HF) and frequency peak of HF-power (fHF) at certain time frame, used for HRVT determination
HF_{RSA}	HF component corresponding to frequencies from 0,04 Hz to the lower limit of locomotion related frequencies
HR	heart rate
HRV	heart rate variability
HRVT	heart rate variability threshold
$HRVT1_{RMSSD}$	first heart rate variability threshold determined from RMSSD values
$HRVT1_{HF}$	first heart rate variability threshold determined from HF*fHF values
$HRVT1_{SD1}$	first heart rate variability threshold determined from SD1 values
$HRVT1_{DFA}$	first heart rate variability threshold determined from DFA-a1 values
$HRVT2_{HF}$	second heart rate variability threshold determined from HF*fHF values
$HRVT2_{SD2}$	second heart rate variability threshold determined from SD2 values
$HRVT2_{DFA}$	second heart rate variability threshold determined from DFA-a1 values
ICC	intraclass correlation coefficient
LF	low frequency (component of HRV power spectrum; 0,04–0,15 Hz)
LOA	limits of agreement
LT	lactate threshold
PNS	parasympathetic nervous system
PSD	power spectral density
RCT	randomized controlled trial (study protocol)
RMSSD	root mean square of successive differences of R–R intervals
RSA	respiratory sinus arrhythmia

SD	standard deviation
SDNN	standard deviation of the NN interval
SNS	sympathetic nervous system
VE	ventilation
VCO_2	carbon dioxide output
VO_2	oxygen uptake
VO_2max	maximal oxygen uptake
VT	ventilatory threshold

CONTENTS

1 INTRODUCTION

Exercising at different intensities causes physiological and biochemical changes in the human body during exercise session. Previous studies have indicated that usually those changes are not occurring linearly when exercise intensity is increased, instead certain turning points can be revealed. The first exercise threshold (ET) presents the work intensity after which there occurs a change in energy metabolism towards anaerobic. Exercise below and at first threshold intensity can be tolerated for multiple hours without increased accumulation of acidity in muscles and blood stream. The second threshold is the highest exercise intensity which is possible to be maintained without constant increase in acidosis. (Nummela & Peltonen 2018, 96-97) Lactate and ventilatory measurements during incremental graded exercise test are traditionally used for training zone boundary determination. The correct determination of training zones is important as even a minimal exceeding of target intensity boundary in training may lead to negative physiological responses and interfere the desired training stimuli. (Gronwald et al. 2020) However, multiple different methods to determine the threshold levels from lactate and ventilatory parameters exist (Faude et al. 2009; Nummela & Peltonen 2018, 79). Thus, differences in the threshold values may occur depending on the method used.

Heart rate variability (HRV) represents the variation of time durations between consecutive heart beats. It is widely used as an indirect method to measure autonomic nervous system (ANS) activity and especially its vagally mediated stimuli. (Acharya et al. 2006) HRV responses to exercise are quite widely studied and accepted. When exercising at low intensities (below 40 % VO_2max) frequency and mathematical variation in heartbeat interval lengths usually start to show a reduction compared to rest. However, at the same time, the complexity of R–R-interval lengths is frequently showing increase as parasympathetic nervous system (PNS) activation is still dominating over sympathetic nervous system (SNS) activity. When exercise intensity rises from low to moderate, ANS modulation changes, and the effect of SNS-activity starts to rise over PNS-activation. Hence HRV, by every measure, decreases time-dependently with increasing intensity. (Hautala et al. 2003) When exercise intensity is increased from moderate to high linear reduction in HRV occurs (Gronwald & Hoos 2019). Changes in ANS-activation are not the only factor affecting HRV whilst exercising. Also, other physiological mechanisms e.g., cardiorespiratory-locomotion coupling, increased filling of the heart and its mechanical effect on sinus node may all independently modify HRV (Cottin et al. 2004; Gronwald & Hoos 2019).

As the behavior of R–R-interval variability and complexity during exercise is widely known, it has been theorized to be appropriate measure for training zone determination, too. Thus, methods for determination of heart rate variability thresholds (HRVT) have been presented in a couple of studies in recent years (Buccheit et al. 2007; Cottin et al. 2005; Cottin et al. 2006; Karapetian et al. 2008; Mateo-March et al. 2022; Mendia-Izueta et al. 2016; Mourot et al. 2014; Rogers et al. 2021a; Rogers et al. 2022; Quinart et al. 2014; Sales et al. 2011). HRVT-determination has been suggested to be cheaper and easier method for the definition of ETs when compared to traditional lactate and ventilatory measurements while it is non-invasive and needs no special test equipment (Rogers et al. 2021a). HRVT can be defined using different HRV-parameters. The determination is based on defining trend changing points (Cottin et al. 2005; Cottin et al. 2006; Karapetian et al. 2008; Mourot et al. 2014; Sales et al. 2011; Quinart et al. 2014) or limit values (Mateo-March et al. 2022; Rogers et al. 2021a; Rogers et al. 2022; Quinart et al. 2014) in the HRV-curve measured during graded exercise test.

HRVTs are reflecting metabolism and homeostasis related mechanisms which are occurring during exercise. The same mechanisms are represented by changes in other physiological variables as well, such as in blood lactate concentration and ventilatory gas exchange. Traditional ventilatory and lactate thresholds are used to capture those mechanisms and because ANS balance is affected by those same changes, it is suggested that modifications in HRV during exercise are representing the same physiological processes (Sales et al. 2011; Tulppo et al. 1996).

Previously, studies examining HRVT determination have been mainly conducted in men and usually in athletic population. In this study all participants were female and had no background in competitive training. The application of HRVT-determination methods into female and non-athletic population is of interest as previous studies have suggested that there might be differences in HRV between sexes. For example, heart rate variability of female has been indicated to be higher than age-matched men's and decreasing slower with age. Also, individual fitness level may affect HRV. (Acharya et al. 2006; Ryan et al. 1994; Tulppo et al. 2001) Hence, the results of this study offer novel information about the applicability of HRVT-methods for training zone determination of novice female runners.

2 HEART RATE VARIABILITY DURING EXERCISE

Heart rate variability (HRV) indicates differences in lengths of consecutive heart beats, and it reflects multiple different physiological factors that modulate cardiac rhythm. HRV analysis enables a practical method for the estimation of autonomic nervous system (ANS) activation status and especially the activity state of its parasympathetic branches. (Acharya et al. 2006) HRV analysis has also applications in clinical settings. Reduced HRV and increased activation of the sympathetic nervous system may reflect increased risk of cardiac complications and arrests (Sandercock et al. 2005a). As it mirrors lowered regulatory capacity to respond to stress factors of ANS it can also correlate with increased mortality risk (Shaffer et al. 2014).

2.1 Physiological background of heart rate variability

As already mentioned, HRV is reflecting changes occurring in time intervals between consecutive heart beats. Those can be determined from ECG-recording by calculating time lengths of successive R-peaks of QRS-complexes (figure 1). Differences in R–R-interval lengths are caused by multiple regulatory systems which target at ensuring the optimal physiological functions and homeostasis. (Shaffer et al. 2014)

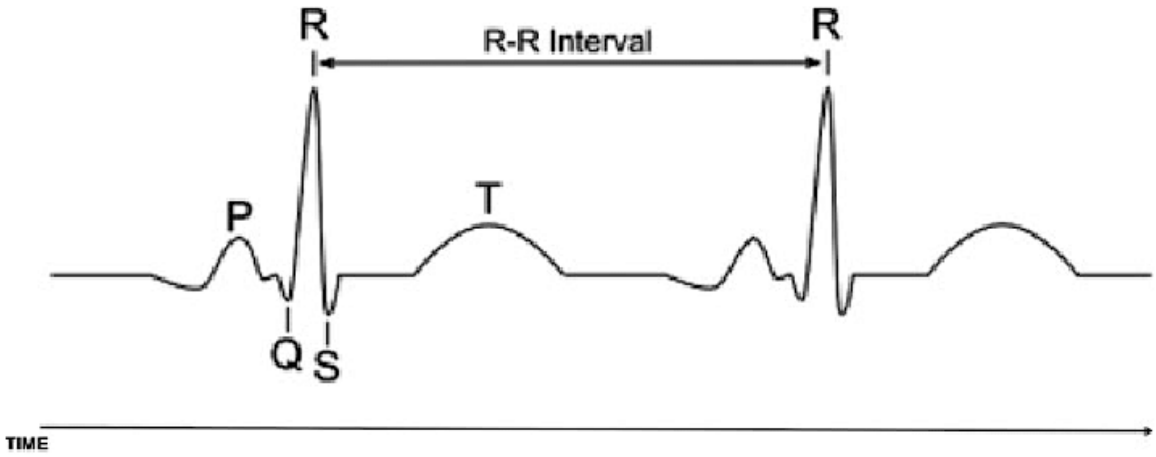


FIGURE 1. R–R-interval lengths for HRV calculations are measured from ECG between consecutive R-peaks (modified from Massaro & Pecchia 2016).

Heart rate variability is used in many practical, research and clinical purposes as a reflector of multiple physiological functions that are modulating cardiac rhythm. HRV analysis enables a

practical method for evaluating the activation level of ANS. ANS is divided into parasympathetic (PNS) and sympathetic nervous system (SNS) and HRV is especially reflecting the activation of parasympathetic branches. (Acharya et al. 2006) Heart rate (HR) decelerates and HRV usually increases when activation of PNS increases (Shaffer et al. 2014). Correspondingly, in situations when the activation of sympathetic branches augments, parasympathetic stimuli diminish, HR accelerates and HRV usually decreases (Shaffer et al. 2014). Thus, there are also nonneural factors affecting HRV which means that HRV never reaches absolute zero level (Cottin et al. 2004). Changes seen in variation between successive heart beats can serve as an indicator of recovery or stress level as well as of cardiac dysfunctions or diseases. (Acharya et al. 2006)

Different individual factors such as age, fitness level and sex (Acharya et al. 2006) as well as the state of functional capacity and chronic comorbidities affect HRV (Karmali et al. 2017). Younger and physically active individuals usually reach highest HRV-values (Acharya et al. 2006). HRV is linearly increasing from childhood to early adulthood until about the age of 20. However, while children have lower total variance than young adults, dynamic measures (complexity and fractal properties) of HRV do not necessarily differ. From adulthood to elderly age the total variance as well as complexity of HRV reduce progressively. (Pikkujämsä et al. 1999) Some sex differences in HRV-levels also exist e.g., HRV of elderly men is usually lowering faster than in females of the same age groups (Acharya et al. 2006). Also, increased high frequency variation and complexity in HR-patterns have been seen in females compared to males (Ryan et al. 1994; Tulppo et al. 2001) while not in all studies (Pikkujämsä et al. 1999). It seems that complexity in HR dynamics decreases with age in both sexes (Ryan et al. 1994).

2.2 The effect of exercise intensity on heart rate variability

Central autonomic network is serving as an anatomical structure that couples different internal and external stimuli together (Gronwald & Hoos 2019). Already quite soon at the onset of exercise there occurs vagal withdrawal (Cottin et al. 2004) which leads to relative change in activity of the two ANS branches (Gronwald & Hoos 2019). This is mainly due to organismic system withdrawal which aims to protect homeodynamic processes causing changes which turns out as increased SNS and decreased PNS activity (Gronwald & Hoos 2019). Even though SNS activity is highly dominating over PNS activity in high exercise intensities it has been

suggested that vagal activation is not completely withdrawn even at maximal intensities (Kannankeril et al. 2004).

When exercising at low to moderate intensities, HR mostly reflects PNS withdrawal. In turn when exercise intensity augments HR mainly represents increased sympathetic activity. (Blasco-Lafarga et al. 2017; Sales et al. 2011; Tulppo et al. 1996) Decrease in parasympathetic tonus followed by increase in sympathetic activity during exercise can be seen as reducing HRV during exercise (Casties et al. 2005; Tulppo et al. 1996). At the same time with gradual increase in HR after the onset of exercise there also occurs change in self-organized regulation of cardiac rhythm. Variability and complexity of R–R-fluctuations first increase until the exercise intensity reach about 40 %VO_2max after which they start to reduce in a time-dependent manner when exercise intensity augments. That may be because after that point SNS activation begins to dominate over PNS activity. (Hautala et al. 2003) Changes in R–R-fluctuations are independent of starting power, increments of intensity and duration of test stages. HRV decreases linearly when changing from moderate to high exercise intensity and has high negative correlations with RPE and VO_2. (Gronwald & Hoos 2019)

While changes in ANS activity are the main source of changes seen in HRV also other physiological functions have an effect on heart rate variability during exercise. For example, activation of complementary neural mechanisms that maintain coordination of locomotor-respiratory coupling during cycling may cause reduction of correlation properties of HRV (Gronwald & Hoos 2019). Thus, in some cases the increased variability seen at the onset of mild exercise may result from increased effect of cardiolocomotor coupling and the "overlapping" effect of rhythmic locomotion on respiratory sinus arrhythmia which is one integral component of HRV (Blain et al. 2009). Exercise in high intensity also results in increased hyperventilation and strengthened filling of the right pump which may induce mechanical effect on the sinus node. (Cottin et al. 2004) It has also been suggested that in a steady-state exercise circulating catecholamines may have even greater role in maintaining tachycardia than the direct neural input into heart (Tulppo et al. 1996). The minor variability in R–R-interval lengths even at maximal intensities is suggested to result from the effect of nonneural factors. For example, increased hyperventilation at the same time with increase in anaerobic metabolism is at least partly responsible of the remaining HRV seen during exercise above VT1 (Cottin et al. 2004).

3 METHODS FOR QUANTIFYING HEART RATE VARIABILITY

Multiple different methods for heart rate variability analysis have been demonstrated in previous studies. In this next section two maybe the most used traditional linear analysis methods, time and frequency domain analysis are presented. In addition, newer non-linear methods are introduced, focusing especially detrended fluctuation analysis and Poincaré plot.

3.1 Linear analysis methods

Linear HRV analysis methods can be divided into time and frequency domain analysis. Time domain analysis is based on statistical measurements of successive R–R-intervals. In frequency domain analysis the spectral distribution of R–R-signal in a certain time frame is calculated. (Buccelletti et al. 2012) The physical and mathematical background and main parameters of those methods are presented in the following sections.

3.1.1 Time domain analysis

Time domain analysis is HRV analysis method which is based on calculation of time interval lengths between successive heart beats from ECG-data (R–R or N–N -intervals). With time domain analysis, it is possible to define e.g., the average R–R-interval length, average HR, and duration of the longest and shortest interval during the measurement phase. (Task Force 1996) Thus, time domain analysis methods are not eligible for depicting dynamic or rhythmic changes in ANS functions that have an influence on HRV, e.g., changes in respiratory frequency (Gronwald & Hoos 2019; Shaffer et al. 2014; Tulppo et al. 1996).

Standard deviation of successive R–R-intervals (SDNN or SD) is one often used HRV parameter, and it represents the variance in interval lengths during measurement. The longer is the duration of analysis phase, the bigger usually is the variance in interval lengths. Correspondingly, when the duration of interval length shortens SD usually diminishes. That is because differences in R–R-interval lengths are most likely slighter in measurement with shorter duration. (Task Force 1996) It is worth remembering that SD is reflecting all factors that affect HRV in a particular time frame together. Hence, it cannot differentiate the influences

of different physiological systems. (Shaffer et al. 2014) SDs are thus comparable only when the durations of measurements are the same (Task Force 1996).

Root mean square of successive differences of R–R intervals (or square root of the mean squared differences of successive N–N-intervals, RMSSD) is another quite widely used parameter of time domain analysis. It correlates with the activation level of the parasympathetic nervous system. RMSSD is a valid parameter for estimating the effect of PNS activity on cardiac functions and short-time changes in HRV. (Goldberger et al. 2006; Task Force 1996)

3.1.2 Frequency domain analysis

Frequency domain analysis of HRV is based on power spectral density (PSD) analysis which illustrates how the power (i.e., variance) of measured signal is divided as a function of frequency (Task Force 1996). The density of HRV signal represents the frequency of consecutive heart beats which practically means the time phase during which certain beating rhythm is occurring. For example, density of 0.1 Hz represents time frame of 10 seconds. (Shaffer et al. 2014) There are some considerations for successful frequency domain analysis. Firstly, sampling rate of at least 250-500 Hz is recommended. If automatic sampling settings are used (e.g., trend or baseline removal) they may have effect on results which is worth considering. Interpolation is a mathematical signal processing procedure, and it is an important step in frequency domain analysis if the data is including any ectopic or missing beats, arrhythmic events or noise that may affect PSD estimations. Interpolation minimizes the error in signal. It can be done either by interpolating erroneous beat on its preceding or successive beats or on its autocorrelated beat. (Task Force 1996)

The power of HRV signal can be divided into frequency components (or bands). Most used of those are high frequency (HF) and low frequency (LF) components. HF-component depicts variability in R–R-intervals occurring usually in densities of 0.15–0.4 Hz (Task Force 1996) but in high intensity exercise it may increase up to 1 Hz (Cottin et al. 2004). Correspondingly, LF-component illustrates variability occurring between 0.04–0.15 Hz. Two more seldom used frequency bands are very low frequency band (VLF: 0.003–0.04 Hz) and ultra-low frequency band (ULF: < 0.003 Hz). (Task Force 1996) Frequency bands are illustrated in figure 2 as a function of power. Frequency components are usually reported either as absolute power units

(ms^2) or normalized units. Normalized units display the relative proportion of a power component in a certain time frame (Task Force 1996).

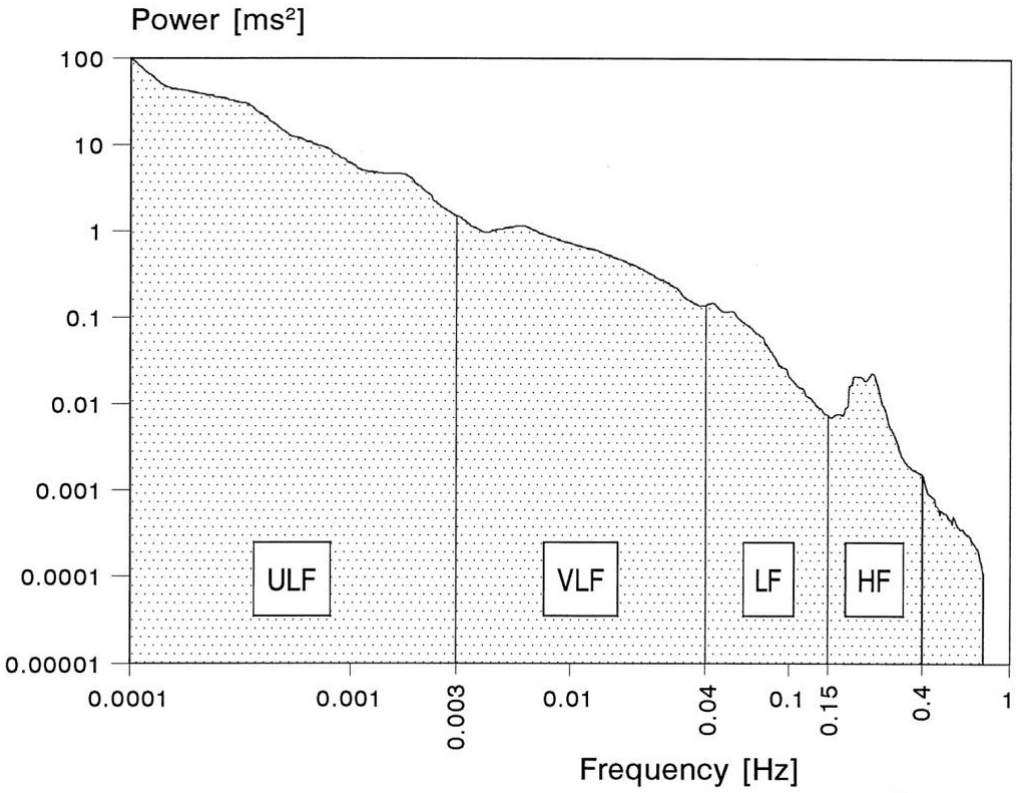


FIGURE 2. The frequency bands of power spectral density of HRV signal (Task Force 1996).

HF-component reflects the functions of the parasympathetic nervous system (i.e., vagal activity) at rest (Cottin et al. 2006; Martinmäki et al. 2006; Task Force 1996) and during exercise (Tulppo et al. 1996). HF-variation is also highly dependent on breathing rhythms. The effect of breathing on HF-variation is called respiratory sinus arrhythmia (RSA). Mechanical effects of breathing cause changes in baroreflex buffering which lead to increased HF-values when exercise intensity increases. That is conflicting with the vagal withdrawal during exercise which has reducing effect on HF-variability per se. (Cottin et al. 2004) Because of that effect, HF-power cannot be used as an unquestionable reflector of parasympathetic functions during exercise.

As the sources of variation seen in HF-power are quite clearly understood, the physiological origin of LF-component is more controversial (Task Force 1996). It seems that the variability of LF-power represents both sympathetic and parasympathetic nerve activity (Cottin et al. 2006;

Task Force 1996) or mainly parasympathetic functions only (Tulppo et al. 1996). LF-variability may also reflect the effect of baroreceptor activation on heart functions. The activation of baroreceptors in blood vessel walls usually leads to decreased blood pressure and SNS activation and increased PNS activation. (Shaffer et al. 2014) The physiological background of VLF and ULF components is equivocal and the analysis of those from short-term recordings is unrecommendable (Task Force 1996). Thus, they are not discussed here in more detail.

3.2 Non-linear analysis methods

In addition to quasi-periodic oscillations, also random fluctuations, and fractal structures (self-similarity) exist in HRV signal. Those can be exposed by using non-linear analysis. (Gronwald et al. 2020) Several neuromuscular, biochemical, central nervous system and peripheral inputs affect heart functions and those are reflected in fractal correlation properties of HRV (Rogers et al. 2021b). Hence, non-linear HRV analysis affords information about the qualitative characteristics of the signal structure and dynamics and possible interactions between physiological subsystems (Gronwald & Hoos 2019). Changes in HR signal correlation properties during diverse organismic demands aim for stabilization and physiological optimization for current or anticipated requirements. That kind of situation is e.g., regulatory withdrawal occurring in organismic regulatory functions when exercise intensity rises. Changes are aiming for systemic integrity and homeostasis. (Rogers et al. 2021b)

Non-linear HRV analysis have been noticed to have some advantages compared to traditional linear analysis methods. One advantage of non-linear methods is that they are adequate for separating time frames which have significantly different HRV but still might have similar spectral and temporal properties (Braun et al. 1998; Task Force 1996). An example of situation of this sort is illustrated in figure 3. Additionally, patterns of HR dynamics which are not detectable by using conventional summary methods of HRV may occur, especially during exercise (Tulppo et al. 1996). The sensitivity of non-linear methods for detecting the effect of different physiological states on HRV is also better when compared to linear analysis methods (Braun et al. 1998). In addition, they are superior for the differentiation of up-going and down-going changes in HR as traditional time and frequency domain parameters do not show information about the direction of time (Braun et al. 1998). Lastly, dynamic HRV measures

may offer complementary information about, for example, aging related cardiac changes (Pikkujämsä et al. 1999).

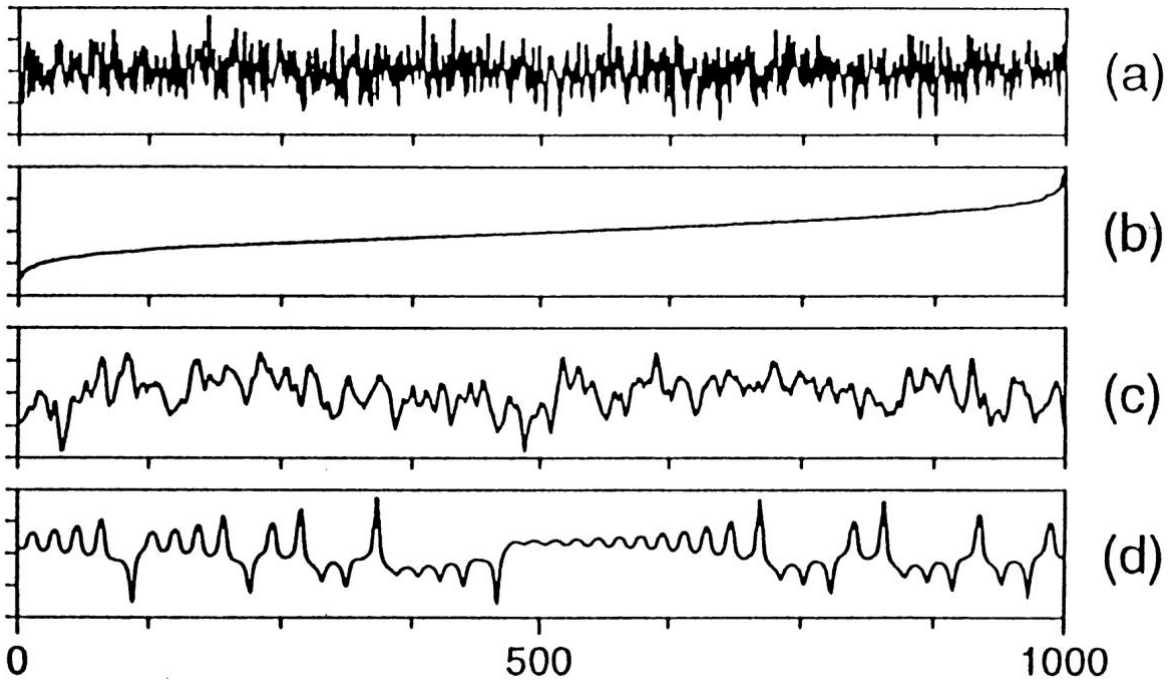


FIGURE 3. An example of four R–R-interval recording series which would have identical HRV if represented by using linear time domain measures only. Series c and d would also have identical power spectrum. (Task Force 1996)

3.2.1 Detrended fluctuation analysis

Detrended Fluctuation Analysis is one of the most used non-linear HRV analysis method. It is useful method for measuring scale-invariant behavior of HRV signal. (Gronwald et al. 2020; Rogers et al. 2021a) The "scale-invariance" is typical for self-affine processes, and it means that scaling of the signal standard deviation is not dependent on the absolute scale. Whereas self-affinity refers to a situation when a certain part of a fractal structure is similar to the whole structure. (Hardstone et al. 2012) Detrended Fluctuation analysis alpha–1 (DFA-a1) is the main parameter of that analysis method (Gronwald et al. 2020; Rogers et al. 2021a). DFA-a1 is suggested to be appropriate parameter for analyzing nonstationary data, e.g., heart beats in a time series and be capable to differentiate physiological demands (Gronwald et al. 2020; Mateo-

March et al. 2022). Nonstationary properties of HRV means that the means, variances and covariances of the signal are changing over time (Gronwald et al. 2020).

Detrended fluctuation analysis is based on cumulative sum of time series and their computation into signal profile. From that profile, set of window sizes is defined and further equally spaced on a logarithmic scale (Hardstone et al. 2012). DFA-a1 is a modification of the root mean square analysis (RMS) and it has low dependence on HR. Analysis of measured HR-signal starts with detrending and integration. After that it is root mean squared. Data is then separated into time windows of different sizes and plotted to a log-log scale against the size of window. (Gronwald et al. 2020; Hautala et al. 2003; Tulppo et al. 2001) The slope of the line can be seen from scaling exponent α which represents the logarithmic relation of data fluctuation to window side (Gronwald & Hoos 2019; Hardstone et al. 2012; Tulppo et al. 2001).

The usage of DFA-a1 method for HRV analysis possesses multiple benefits. One of the strengths of that method is its dimensionless properties. Because of those, there is no need for calibration to other internal load parameters to the same degree than when gas exchange or lactate values are investigated (Rogers et al. 2021b). Another benefit is that DFA-a1 analysis enables comparisons between HRV recordings with different lengths. Additionally, it is accurate for detecting short-term changes in HRV that are caused by changes in nervous system activity but sometimes erroneously associated only to respiration (Blasco-Lafarga et al. 2017). Correspondingly, DFA-a1 can be used for detection of long-range correlations within a time frame while avoiding the erroneous detection of long-range correlations which are caused by nonstationary properties of the signal (Mateo-March et al. 2022). DFA-a1 method is also suited for analysis of relatively short time phases (Gronwald & Hoos 2019; Tulppo et al. 2001). Required recording time window width is only 4-16 beats and for valid calculation about 200 beats or 2-3 min time frame (depending on bpm) is needed. In addition, reaching of steady state is not mandatory. Thus, DFA-a1 method is suggested to be suitable and applicable for different sport-specific settings e.g., for exercise from low to high intensities. (Gronwald et al. 2020; Hautala et al. 2003) DFA-a1 also seems to be “robust” against artefacts (Gronwald & Hoos 2019) and its test-retest reliability is high (Hardstone et al. 2012).

When HRV is presented by DFA-a1, the values are usually ranging from 0.5 to 1.5. DFA-a1 value 1.5 reflects strongly correlated HR pattern, value 1.0 mix of uncorrelated and maximally correlated beats and values below 0.5 anti-correlated pattern (Hautala et al. 2003). When DFA-

a1 is between 0.5-1 positive correlations and larger fluctuations than what would be expected by chance are occurring on longer timescales. Correspondingly, if DFA-a1 reaches value below 0.5 time series is anti-correlated, and fluctuations are smaller in larger windows than expected by chance. (Hardstone et al. 2012) DFA-a1 values below 0.5 reflect the activation of protective feedback system and stabilizing mechanisms which are activated at the same time when interactions and coordination between different subsystems stop working (before the function of the whole system termines) (Rogers et al. 2021b). That mechanism aims to immediately correct fluctuations to the opposite direction (Gronwald & Hoos 2019). The scale of DFA-a1 values is represented in figure 4.

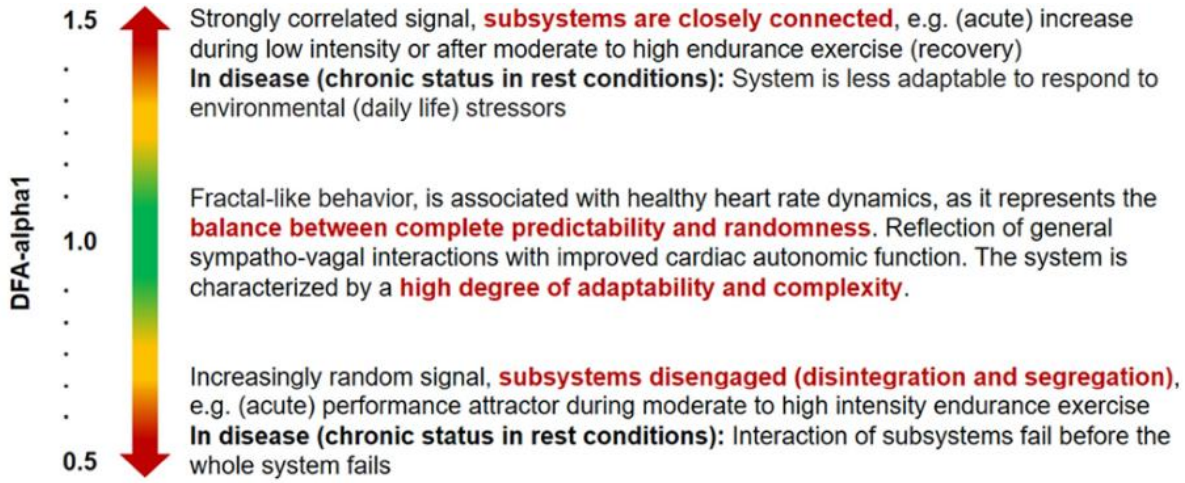


FIGURE 4. The scale of DFA-a1 values during different physiological situations (Gronwald et al. 2020).

DFA-a1 is dimensionless by nature and it is readily "calibrated" to external load because of the expression of fractal patterns of HR due to internal physiological changes (Rogers et al. 2021c). DFA-a1 method has been suggested to be suitable for different applications. It may be practical for constant updating of training intensity zones and thus optimalization of training (Mateo-March et al. 2022). The method has also been used for assessing cardiovascular risk in clinical settings (rest measurement). There is evidence that DFA-a1 values differing from normal 1.0 may reflect higher morbidity or worse prognosis risk. (Gronwald et al. 2020)

Even though DFA-a1 method has clear benefits there are some issues to be considered when analyzing results. Individual fitness level affects DFA-a1 values obtained. DFA-a1 values of individual with lower training status lower at much lower intensity than those of individual with

higher fitness level (Gronwald et al. 2020). In addition to training status, spontaneous breathing frequency may affect DFA-a1 during exercise by enabling physiologic coupling processes (Gronwald et al. 2020). Measurement duration may also have an influence on DFA-a1 values even though data regarding that is weak (Gronwald & Hoos 2019). All in all, DFA-a1 is an easy-to-access and adequate measure for assessing acute and possibly chronic adaptational exercise responses, as well, in a differentiated and qualitative manner. It seems also suitable for diagnostic, monitoring and training control use combined with other internal and external load measures. However, further studies are needed considering the reliability of DFA-a1 during especially high exercise intensities, in both sexes and in people of greater age-range. (Gronwald & Hoos 2019)

3.2.2 Poincaré plot -analysis

Poincaré plot -analysis method enables calculation of changes in heart dynamics with trends. It is based on a scattergram to which each individual R–R-interval is plotted as a function of the preceding interval length. (Tulppo et al. 1996) From that plot diagram two standard deviations, SD1 and SD2, can be determined. SD1 line splits the diagram latitudinally (Tulppo et al. 1998) and it reflects the regulatory effect of PNS on HR without the confounding effect of trends in HR itself (Nascimento et al. 2017; Tulppo et al. 1996). SD2 traverses the diagram longitudinally and perpendicularly in relation to SD1 (Tulppo et al. 1998). SD2 is more dependent on sympathetic activation, too, and its values correlate negatively with increased sympathetic activation and exercise intensity (Tulppo et al. 1996). In conclusion, SD1 reflects changes in R–R-interval lengths in short duration and SD2 in longer duration (Nascimento et al. 2017).

It has been suggested that non-linear Poincaré plot -analysis could be a better predictor of exercise related, acute and longterm changes in HRV than traditional linear methods (Mourot et al. 2004; Tulppo et al. 1996). During exercise, various physiological cardiac changes occur which may be difficult to capture by linear HRV analysis only. (Tulppo et al. 1996) The strength of Poincaré plot is that stationary signal is not needed (Mourot et al. 2004). During exercise, especially parasympathetic activity changes may be better reflected by Poincaré plot (Tulppo et al. 1996).

4 EXERCISE INTENSITY THRESHOLDS

Graded maximal exercise test is recommended for testing endurance performance and aerobic metabolism during exercise, and it is used for assessing exercise threshold (ET) levels. The starting speed or power of graded exercise test (GXT) is guided to set to an individually easy effort level and the effort is increased graded until testee reach her maximum. (Nummela & Peltonen 2018, 80) If the aim is to determine steady state effort level at a certain work intensity (speed/power) it is recommended to use at least 2 min stages in the test, preferably 3 min. Thus, there is enough time for physiological markers to reach steady level. (Nummela & Peltonen 2018, 79)

4.1 Factors affecting endurance capacity and performance

Endurance refers to the ability of muscles to tolerate and produce submaximal forces repeatedly (Kenney et al. 2015, 225). Endurance performance during long-lasting exercise session is basically dependent on oxygen supply from inhaled air to active muscle tissue and the ability of muscle tissue to use oxygen. Therefore, individual's endurance capacity is mainly affected and tested by VO_2max and ET levels. More seldomly tested factors that also have effect on endurance properties include economy of movement patterns, electrical properties of activated muscles, force generating abilities of neuromotor system and the ability of tissues to sustain ischemia. (Peltonen & Nummela 2018, 65) Thus, targets of endurance training are to improve body's ability for oxygen distribution and energy production by enhancing function of lungs, circulatory system, oxidative enzymes, mitochondrial functions, and anaerobic power. In addition, improvements in fatigue tolerance and metabolism by enhanced carbohydrate stores, fat oxidation, buffering capacity, neural functions and lactate removal are desired results of endurance training. (Peltonen & Nummela 2018, 64) In a cellular level endurance training results in multiple adaptations in phenotype of skeletal muscles. Those adaptations include fiber type transformation from glycolytic to oxidative, increased fat use in energy production, angiogenesis, and enhanced glucose uptake into muscles. They are happening via different molecular mechanisms e.g., via increased expression of PGC-1α that is a signal molecule involved in biogenesis of mitochondria. (Kenney et al. 2015, 275)

A couple of factors which affect individual capacity to improve endurance properties exist. One of those is training status. The higher is individual fitness level, the bigger training volume is needed for reaching improvements in endurance performance. (Kenney et al. 2015, 281) Furthermore, the fitter the individual is the lower is blood lactate at certain work rate. Also, patients that are suffering from some cardiovascular disease usually have higher lactate levels during exercise. (Wasserman et al. 1973) Also, genetic factors influence the limits of trainability (Kenney et al. 2015, 282). The effect of sex on capacity to improve endurance is conflicting but possibly minor. Endurance capacity might be slightly better improvable in males than females because of endurance training when measured by changes in VO_2max (Diaz-Canestro & Montero 2019) while sex differences in improvement of VO_{2peak} are not proved (Hautala et al. 2006). Untrained males usually have higher VO_2max than untrained females but in trained people the difference is minor (Kenney et al. 2015, 282). Work effort and oxygen uptake at first threshold intensity are usually higher in men and younger people but there seems to be no differences when values are corrected with weight (Reinhard et al. 1979).

The testing circumstances and protocol may also have an influence on measured endurance variables. GXT stage duration and load increments can influence cardiorespiratory and metabolic responses. Longer-lasting protocols (20-30 min with increments of 3-5 min) may prevent the subject from reaching absolute maximal aerobic speed because of accumulation of dehydration, acidosis in muscles, fatigue, and cardiovascular drift. In cycling for example, shorter GXTs (12-14 min) with 1 min stages have turned out to be valid for the estimation of VTs and LTs as well as VO_2max. (Cerezuela-Espejo et al. 2018) Time which is required for VO_2 to reach steady state is related to work intensity. The more intense is the effort level the longer time it takes to reach steady state. In moderate effort exercise steady state in VO_2 is reached within 4 min but it can take even 10 min or longer when intensity is augmented to very heavy and in extreme cases it could not be reached at all. (Wasserman et al. 1967) In Finland, 3 min stages in GXTs are recommended when assessing thresholds (Nummela & Peltonen 2018, 81).

4.2 Traditional methods for threshold determination

In many studies the first ET is called aerobic threshold (AerT), and it is useful parameter for defining the suitable exercise intensity for healthy weight maintenance and for developing physical fitness (Zimatore et al. 2021). It is also applicable for clinical use when assessing exercise tolerance of patients with cardiorespiratory disease (Wasserman & McIlroy 1964). Work intensity which is below first threshold can be sustained for hours. The physiological changes occurring at the intensity level of first threshold include the first increase in blood lactate levels, increased venous H^+ and K^+ concentration and activation of chemoreceptors because of fluctuations in PO_2 and PCO_2 levels. If the work intensity is yet raised, the release of lactate into blood is linearly increased and the production and ventilation of CO_2 are further augmented in relation to oxygen consumption. In many studies, the second threshold is called anaerobic threshold (AnaT) which refers to the intensity above which ventilation begins to accelerate more in relation to CO_2 production to compensate the increased acidosis. This is also the highest work intensity which can be sustained without prolonged increase in lactate concentration. (Nummela & Peltonen 2018, 96-97)

Physiological ETs can be determined by different ways. Critical power (Vanhatalo et al. 2007), maximal lactate steady state (MLSS; Baron et al. 2008), respiratory compensation point (Bergström et al. 2013), onset of blood lactate accumulation (OBLA; Sjödin & Jacobs 1981) and deoxyhemoglobin breakpoint (Turnes et al. 2019) are methods which have been used for thresholds assessment in previous studies. Physiological background of these physiological markers is similar because they are all reflecting same cellular metabolism and chemical changes inside the body which are occurring with increasing work effort. Already years ago, it has been suggested that lactate and VO_2 would reach steady state at about the same exercise intensity. (Wasserman et al. 1967) By date, the evidence about the relationship between lactate thresholds (LT) and ventilatory thresholds (VT) is controversial (Cerezuela-Espejo et al. 2018). That is why, for example, in Finland LTs and VTs are nowadays recommended to be reported independently from each other (Nummela & Peltonen 2018, 96) even though it has been suggested that values of physiological parameters (VO_2, HR and blood lactate) in those threshold intensities are behaving similarly in young, fit people (Ahmaidi et al. 1992) and in untrained adult men (Aunola & Rusko 1984). Recent studies have shown that metabolic pathways behind for example MLSS and VT2 are different and MLSS settles more likely at the

midpoint of VTs, although, it has previously been suggested that VT2 would correlate with MLSS (Cerezuela-Espejo et al. 2018).

Because there may be detectable differences in threshold levels assessed by different physiological measures, the use of same parameter when analyzing training program efficiency is important (Nummela & Peltonen 2018, 79). Thresholds determined from lactate or ventilatory gas measures do not always agree that might result in inaccurate training zone boundaries. Furthermore, differences between distinct tests of the same individual may exist even though the method used would be the same. That may be problematic in practical training as even minimal exceeding of target breakpoint may result in glycogen depletion, prolonged parasympathetic cardiac recovery, gastrointestinal disruptions and increased overall central and muscular fatigue (Gronwald et al. 2020) that may weaken the desired training stimuli. Another consideration with the practical use of ETs is that both LTs and VTs are strongly dependent on test protocol used. Thus, thresholds are not necessarily comparable to constant load exercise intensities. Also differences in biomechanical efficiency between laboratory treadmills or ergometries vs. trail running, road cycling or skiing may affect the applicability of results obtained for practical use. (Gronwald et al. 2020)

4.2.1 Lactate thresholds

Determination of lactate thresholds (LT) is taken as a criterion measure for aerobic endurance capacity as LTs have been shown to be well correlated with competition performance in running and less strongly also in cycling. Hence, studies in other endurance sports are quite rare. (Faude et al. 2009) Increased arterial lactate concentration during exercise reflects the increased production of lactate when compared to its catabolism which is resulting from numerous different physiological and biochemical changes, e.g., increased need for ATP production via glycolysis (Gladden 2004; Robergs et al. 2004). Physiological functions which are occurring besides lactate accumulation into blood include impairments in metabolic acidosis and muscle contraction, altered oxygen kinetics and increased hyperventilation. As a result of those, exercise capacity deteriorates. (Myers & Ashley 1997)

During GXT blood lactate levels increase exponentially with intensity increments. Changes in blood lactate are minor when work intensity is moderate (Wasserman et al 1967). Oxygen

dependent energy metabolism is the most economical way to produce ATP and lactate release to blood is slight. LT1 refers to the first increase in blood lactate concentration from baseline levels. That occurs because of increased proportion of anaerobic energy production. When intensity is increased above LT1, lactate production as a metabolic byproduct augments and blood lactate level increases linearly at the same time with intensity increments. LT2 is the highest exercise intensity which can be sustained without constant rise in blood lactate concentration, occurring usually near maximal lactate steady state (MLSS). (Nummela & Peltonen 2018, 97) At that point, lactate production and removal are still equal (Nummela & Peltonen 2018, 97) and lactate concentration reaches plateau (Wasserman et al. 1967). If exercise intensity is further increased from that, lactate concentration increases exponentially until it reaches its peak value (Wasserman et al. 1967).

Multiple concepts exist for assessing ETs according to blood lactate concentration curve. Those include the use of fixed blood lactate value, the detection of the first rise in lactate levels above baseline and different "threshold concepts" which are based on, for example, detection of MLSS or rapid changes in lactate curve. In fixed blood lactate concepts, LT is usually set to a point at which lactate concentration reaches value of 2, 2.5, 3 or 4 mmol/l. In threshold concepts LT is set to the intensity stage at which there occurs first significant/non-linear/sharp/marked/abrupt increase above baseline in blood lactate concentration curve. (Faude et al. 2009) In Finland, it is agreed to determine LT1 to the GXT intensity level at which blood lactate for the first time reaches value of +0.3 mmol/l above the lowest lactate concentration measured during test. Finnish recommendation for LT2 determination suggests the fitting of two linear lines into the lactate curve. One line is drawn through LT1 and the following test stage and other through the lactate values of the last test stages between which the increase in blood lactate is above 0.8 mmol/l. LT2 is defined as test stage intensity which is interposed to the intersection of those two lines. (Nummela & Peltonen 2018, 97) The Finnish guidelines for LT determination are illustrated in figure 5.

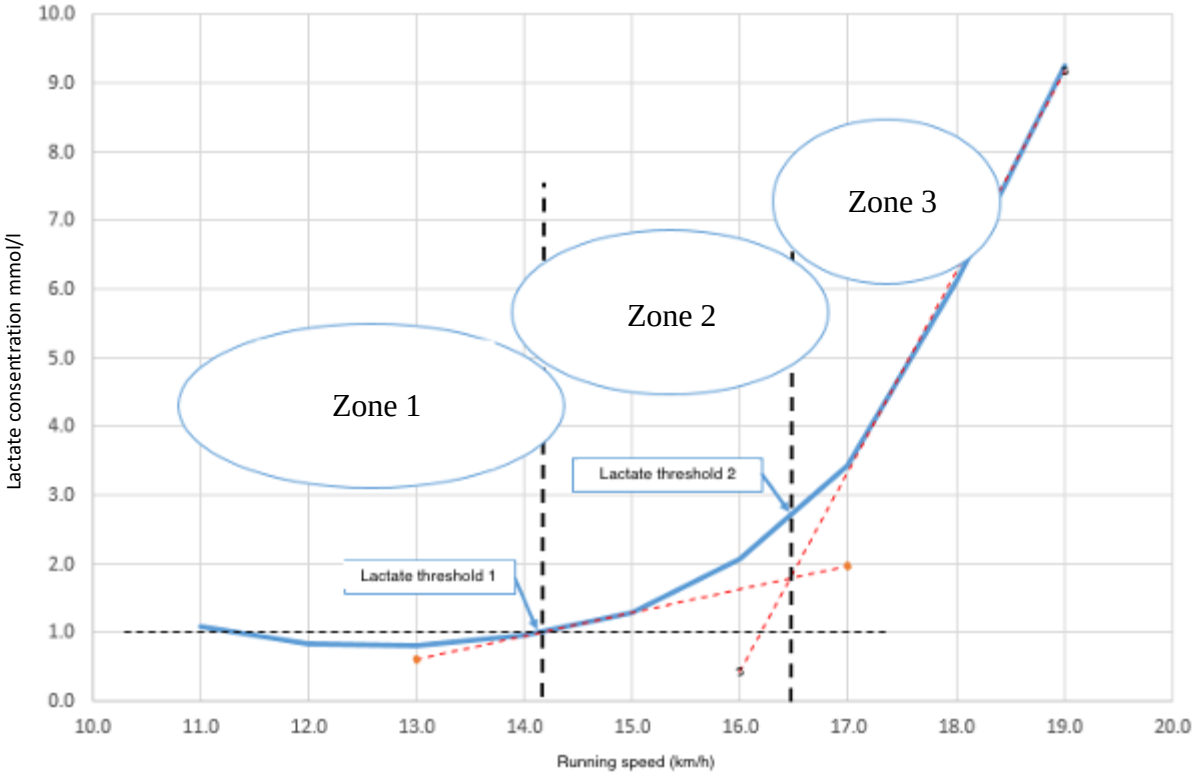


FIGURE 5. Determination of lactate thresholds according to Finnish guidelines (modified from Nummela 2021).

The reliability of lactate threshold concepts has basically been considered high and the between-measures variation low (Pfitzinger & Freeson 1998). There are yet some considerations to be considered. Firstly, LT1 may be difficult to assess as clearly detectable changes in lactate levels seldom occur during the first stages of GXT. Also lactate analyzers might feature typical error during low intensity zones. To take this possible bias in lactate values at low intensities into account, LT1 has been determined as the workload 0.2 mmol/l above the lowest value measured during test. Also, +0.5 mmol/l increase from resting value is used as well as the lowest value from calculation where lactate concentration is divided by work rate or VO_2 (=minimum lactate equivalent). (Faude et al. 2009) In Finnish guidelines the 0.3 mmol/l increase from lowest measured lactate concentration is recommended (Nummela & Peltonen 2018, 97). However, day to day variation in resting lactate concentration exist to some degree which may affect to LT determination methods in which resting levels are taken into account (Pfitzinger & Freedson 1998). Methodological diversity is also evident when visual LT1 determination methods have been compared (for example log-log and linear spline intersection methods). Also, threshold values used for LT1 differ as in some studies 1 mmol/l lactate rise from baseline is used while in others fixed value of 2 mmol/l. (Gronwald et al. 2020) In one study conducted on a treadmill,

LT1 was determined as the highest speed which induces no detectable rise in blood lactate values above baseline. The threshold was on average 0.8 mmol/l above resting values and correlated with VT1. (Cerezuela-Espejo et al. 2018) Because of interpersonal differences in lactate accumulation, the analysis of relative changes in lactate concentrations during GXT may be more favorable compared to absolute lactate values. In addition to those methodological issues, visual determination between different researchers may cause variation to results which should be taken into account. (Faude et al. 2009)

For LT2 determination multiple definitions and methods are used, as well. Examples of those are the point at which fixed value of 4 mmol/l is reached, maximal lactate steady state (MLSS) and various computer model-based calculations. Different definitions are usually resulting in distinct values. In addition to method used, results may also depend on test protocol and expertise for visual interpretation of the researcher. (Rogers et al 2021b) Fixed value of 4 mmol/l is used in some studies and it is explained to be the highest lactate concentration which is sustainable in longer term (Faude et al. 2009). However, considerable differences between (Faude et al. 2009) and within (Aunola & Rusko 1984) individuals have been detected. For example, it has been shown that women reach a lactate value of 2 mmol/l at higher relative intensities than men while the log-log assessed LT1 is comparative. (Gronwald et al. 2020) A study by Cerezuela-Espejo et al. (2018) suggested that LT1 +3.0 mmol/l best predicted VT2 when different LT2 methods were compared. However, blood lactate values may differ between measurements of same individual due to, for example, factors relating carbohydrate and fat metabolism (Aunola & Rusko 1984). MLSS have been suggested to correlate with the transition point between VT1 and VT2. The problem with the use of that parameter is that possible differences exist in its values depending on test mode. For example, the energy expenditure in running at MLSS intensity may be higher than in cycling (Cerezuela-Espejo et al. 2018).

The use of individual anaerobic threshold (IAT), that is determined as an intensity at which inclination in the lactate curve exist, is nowadays recommended (Faude et al. 2009). In one application of that method two tangents from the lower rates and from the higher rates of GXT are determined and IAT defined as the intensity of the stage at their intersection (Faude et al. 2009) as also Finnish guidelines are suggesting (Nummela & Peltonen 2018, 97). Dmax method is also quite widely used for LT2 assessment. It is based on detecting the maximal perpendicular distance of the curve when line is drawn connecting the starting and end points of the GXT lactate curve (Faude et al. 2009). In some studies, LT2 is also determined by visually inspecting

the last point in the lactate curve before the second non-linear and exponential increase in blood lactate occurs (Ahmaidi et al. 1992). It is still worth remembering that all these methods depend on starting intensity and on how maximal is the effort of testees (Faude et al. 2009).

Even if LT assessment procedure would be conducted carefully according to chosen standards factors that may affect lactate values exist. Firstly, differences in test stage duration or magnitude of increases in work rate may affect lactate curve and threshold values (Gronwald et al. 2020). In addition, the sampling method may influence results, too. Blood samples from earlobe usually result in lower lactate values than samples from finger. Blood sampling from plasma results in considerable higher lactate values than whole venous samples. Samples taken by capillary method settle in between those. Some studies have shown even considerable differences between lactate values measured by using different lactate analyzers (portable field analyzers, laboratory analyzers) and methods (amperometric vs photometric methods). Also, climatic conditions may influence blood lactate concentration. (Faude et al. 2009) The effect of diet on blood lactate levels measured during GXT is controversial. In some studies, they have suggested that diet has no effect on lactate threshold values obtained (Quirion et al. 1988; Yoshida et al. 1984) while in others, e.g., reduced iron concentration in blood is found to increase blood lactate levels (Ohira et al. 1981). All in all, for reliable LT determination used methods, sampling procedures and test preceding circumstances should be standardized as carefully as possible.

4.2.2 Ventilatory thresholds

In addition to blood lactate, measurement of various ventilatory parameters is another quite widely used method for ET assessment. At the onset of mild exercise, ventilation in relation to oxygen consumption (VE/VO_2=ventilation equivalent for oxygen) begins to decrease because of improved gas exchange in the tissues. It decreases until it reaches its minimum at about one third of maximal exercise capacity and then starts to increase. (Reinhard et al. 1979) Ventilation per minute (VE) is increasing in proportion to oxygen consumption until certain intensity point is reached after which VE begins to increase more rapidly when compared to VO_2 (Reinhard et al. 1979; Wasserman et al. 1973). Same disproportional increase occurs also in VE/VCO_2 but it is usually less clear than when VE is compared to VO_2 (Wasserman et al. 1967). Bicarbonate is a substance that functions as an acid buffer in the blood stream. It acts for keeping pH at a

normal level by increasing CO_2 production for increased metabolic acidosis to be compensated. (Wasserman et al. 1967; Wasserman et al. 1973) First ventilatory threshold (VT1) results from hyperpnea which is provoked by increased CO_2 production from metabolism when exercise intensity rises. At that point VE/VO_2 relationship and end-tidal oxygen pressure ($PetO_2$) increase nonlinearly while VE/VCO_2 and end-tidal carbon dioxide pressure ($PetCO_2$) remains constant (Reinhard et al. 1979; Wasserman et al. 1973). Chemoreceptors of the peripheral arteries forward the information about changed ion concentrations via neural networks and in response to that ventilation is increased (Nummela & Peltonen 2018, 96). VE is important and sensitive part of respiratory control mechanisms for the elimination of increased CO_2 from buffering reactions without changes in $PetCO_2$ (Wasserman et al. 1973).

When exercise intensity is further increased CO_2 production and ventilation are accelerating in relation to oxygen consumption (Reinhard et al. 1979; Wasserman et al. 1973). Phase between VT1 and VT2 is also called isocapnic which means that ventilation and CO_2 production are increased equally (Nummela & Peltonen 2018, 96). The compensation of acidosis is possible to a certain intensity degree by increasing ventilation (Wasserman et al. 1967). However, at some point ventilation begins to accelerate faster in relation to VCO_2 (Ahmaidi et al. 1992; Wasserman et al. 1973). VT2 is reached when the compensation of increased CO_2 production by hyperpnea is not sufficient anymore and results in increase of VE/VCO_2 (Ahmaidi et al. 1992; Reinhard et al. 1979) which can be seen as a first systematic and non-linear increase in VCO_2 in heavy intensity exercise (Ahmaidi et al. 1992). At that point reduction of bicarbonate becomes more distinguishable and the stimulation of ventilation increases and leads to decrease in $PetCO_2$ (Wasserman et al. 1973).

5 HEART RATE VARIABILITY THRESHOLDS

The purpose for exercise thresholds (ETs) is to capture mechanisms related to homeostasis and metabolism which are occurring during exercise. The reason for the activation of those mechanisms is physiological changes explained in previous sections. Those mechanisms are related to both LT and VT and ANS balance, thus, it can be assumed that changes seen in heart rate variability (HRV) reflect the same breakout points (Sales et al. 2011). There is evidence that observable changes in autonomic nervous system status and ventilatory parameters are related (Tulppo et al. 1996).

Traditional "gold standard"-methods for assessing thresholds are usually costly and require experienced users, regular calibration, and special testing equipment (Rogers et al 2021a). Hence, heart rate variability threshold (HRVT) methods are suggested to be cost-effective, easy availability and non-invasive options for assessing ETs for healthy individuals (Rogers et al. 2022) with different physical condition (Zimatore et al. 2021) as well as for clinical population e.g., diabetes patients (Sales et al. 2011). Nowadays, the monitoring of HR by non-invasive and cheap mobile applications is easy, thus, HRVT determination may be applicable method for field use (Zimatore et al. 2021). Additionally, HRVTs have shown good reproducibility, low bias (Nascimento et al. 2017; Novelli et al. 2018) and low artefact percent (< 5 %) even when different HR monitoring devices have been used and compared (Rogers et al. 2022).

5.1 Assessment of first heart rate variability threshold

Multiple different HRV methods exist and have been used for first ET assessment. Some of them are presented in next sections.

Time domain analysis. The first ET has been determined in a couple of previous studies by using RMSSD-parameter of time domain analysis. HRVT1 has been set to the point in RMSSD curve at which the deflection of the curve exists or at which values are lowered below 3 ms^2. The threshold has been set to a GXT stage after which there occurs no more decreasing trend, despite intensity increments. The RMSSD-values averaged from the last 60–120 s of each stage are reported and plotted against workload. (Karapetian et al. 2008; Mendia-Izueta et al. 2016; Novelli et al. 2018; Queiroz et al. 2017; Quinart et al. 2014) HRVT determination may be partly

problematic by using time domain parameters, especially in sports where upper body movements are needed, e.g., in skiing (Mendia-Iztueta et al. 2016). RMSSD is also sensitive to load and cadence changes during exercise. Trend changing points in RMSSD curve which are seen at about the first threshold effort level are also suggested to reflect command anticipatory response to exercise or increased strength levels. (Blasco-Lafarga et al. 2017) Thus, it is not clear if they are purely reflecting the physiological changes associated with first threshold intensity.

Frequency domain analysis. When considering the power spectral density of HRV during exercise, low frequency (LF) fluctuations are predominating over high frequency (HF) fluctuations when exercise intensity is moderate, while the opposite situation occurs when intensity increases above VT1 (Casties et al. 2005; Cottin et al. 2004). Significant decrease in HRV spectral energy is visible when exercise above VT1 work rate is compared to that of below. The reason behind that effect are qualitative changes in cardiac control that are occurring when exercise intensity is increased. High frequency power of HRV reaches its minimal value just before the first threshold as there occurs concomitant increase in tidal volume and BF at the same time when vagal activity withdraws. Those may be resulting from mechanical effect on the sinus node by hyperpnea. (Cottin et al. 2005; Cottin et al. 2006) Correspondingly, increased HF over LF during very heavy exercise may be due to increased BF combined with reduced autonomic control of the heart (Cottin et al. 2004).

Cottin et al. (2005 & 2006) investigated the correlation between VTs and HRVTs determined by two different methods: by using frequency peak of HF-power (fHF) for each 20 s period during GXT and by using mathematical product of fHF and high frequency power (HF) averaged similarly for each 20 s period. Both values were then plotted on a curve against time and work rate. Quinart et al. (2014) used Fourier transform with 64 s sliding window with recalculation every 3 s. In all three aforementioned studies, first threshold (HRVT1) was determined visually from the curves of both variables as the intensity stage in which first nonlinear increase occurred. (Cottin et al. 2005; Cottin et al. 2006; Quinart et al. 2014) However, the problem with threshold assessment from HF-component might be that the reduction and stabilization points in its values are not easily detectable in all individuals. (Cottin et al. 2006) The changes in variability measured by frequency domain parameters may be very slight during acute exercise bout making the detection of trend changing points difficult (Hautala et al. 2003). On the other hand, spectral components of HRV may better represent VT

than time domain parameters which purely represent the changes in ANS activity. That is because according to current knowledge ANS is not directly affecting ventilatory parameters according to which VTs are determined. (Quinart et al. 2014) Spectral analysis can also be conducted from time phases of shorter duration than which are required for time domain analysis. The accuracy of threshold HR and intensity values would then be superior. (Quinart et al. 2014)

Poincaré plot. Non-linear methods of HRV may serve as a practical reflector of heart behavior during especially high intensity exercise when traditional HRV analysis methods seem to be insufficient (Casties et al. 2005). Poincaré plot is one non-linear method which has been used for the determination of HRVT1. SD1 parameter of that method has been shown to correlate with PNS activity (Tulppo et al. 1996). HRVT1 identification procedures which are used in previous studies are similar to RMSSD-based methods. First threshold has been identified to a workload during which first breakpoint in SD1 curve occurs (Cambri et al. 2016; Nascimento et al. 2017; Queiroz et al. 2017), during which the values are stabilized below 3 ms^2 (Novelli et al. 2018; Sales et al. 2011; Queiroz et al. 2017) or to the first workload in which difference in SD1-value is below 1 ms^2 when compared with SD1-value of the preceding workload (Cambri et al. 2016). In addition to visual methods, also Dmax-method has been used. It is based on determining a linear regression through the first and last value of SD1 or SD2 curve. From that curve the longest perpendicular distance to the regression line is defined as threshold. (Nascimento et al. 2019) The trend changing point in SD1-curve is usually seen at about 60% of maximal effort level and it has been suggested to reflect same physiological functions as RMSSD. However, while RMSSD is more sensitive to the changes in load and cadence, SD1 is more sensitive to brain function and psychological changes. (Blasco-Lafarga et al. 2017).

Detrended fluctuation analysis. Traditional HRV methods are usually at least partly depending on sinus node firing rate. Instead, fractal methods are not affected by pacemaker firing and theoretically reflect both SNS and PNS activity. (Blasco-Lafarga et al. 2017) Recently, DFA-a1 of detrended fluctuation analysis have been tested and noticed to be potential variable for threshold determination. During graded exercise DFA-a1 behaves biphasic; when exercise intensity is very low to moderate values are same than at rest (1.0-1.5) or even increased (up to 1.5) until the intensity reaches 40 %VO_2max (Hautala et al. 2003). Prolonged, low intensity exercise reintegrates and synchronizes subsystems which leads to increased DFA-a1 (Gronwald et al. 2006; Hautala et al. 2003; Tulppo et al. 2001) while in prolonged exercise of low-to-

moderate intensity (around first threshold level and above that) the onset of disintegrating processes occurs (Gronwald et al. 2006; Hautala et al. 2003). After that, DFA-a1 values start to decrease (Hautala et al. 2003) and the decrease is significantly speeded up when exercise intensity is augmented over 60% of maximal, the effort level which is often associated with first ET. Already at that point, PNS activity is supposed to be almost entirely inhibited. (Blasco-Lafarga et al. 2017) Thus, HRVT1 has been determined to the time point in which DFA-a1 reaches value of 0.75 as it represents the midline between fractal behavior of HRV (1.0) and an uncorrelated, noisy behavior (0.5). In previous studies DFA-a1 values have been calculated using sliding 2 min time window with recalculation every 5 seconds during GXT. When plotted to time, threshold VO_2, HR and effort level could be defined. (Rogers et al. 2021a; Rogers et al. 2022; Mateo-March et al. 2022) An example of HRVT1 determination using detrended fluctuation analysis is presented in figure 6.

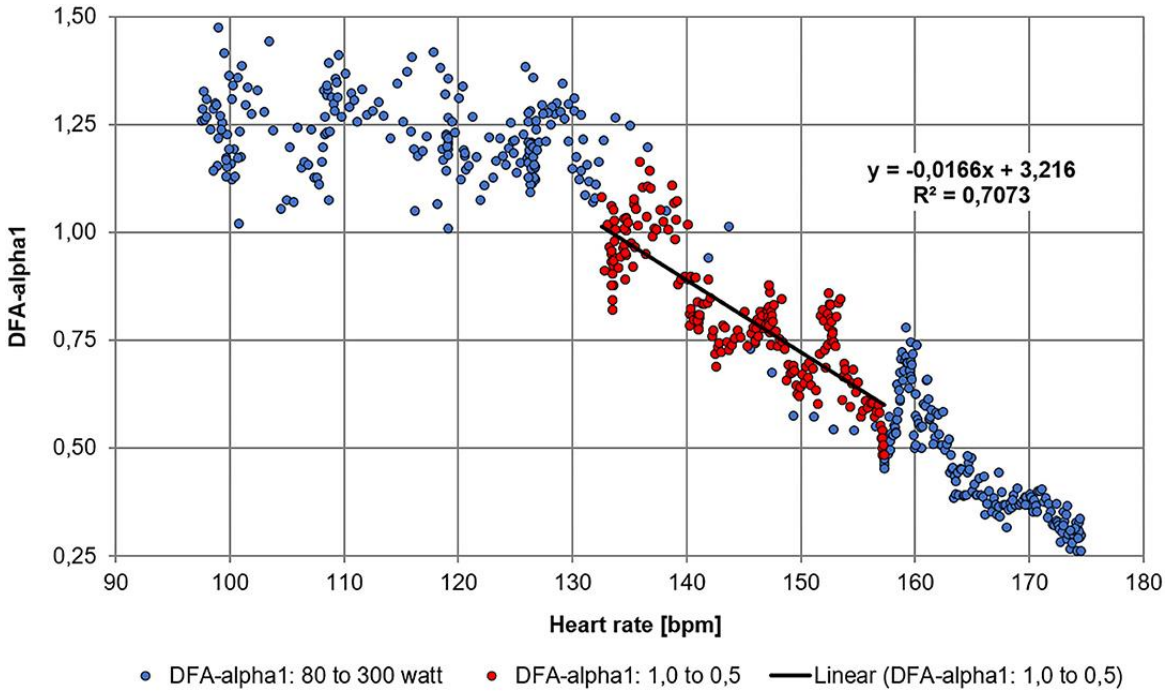


FIGURE 6. An example of HRVT determination from DFA-a1 values measured during graded exercise test (Gronwald et al. 2021).

Compared to other HRV parameters proposed for definition of first ET, the benefit of DFA-a1 is that high intensity zones are not needed thus it may be potential in defining exercise zones in populations that are not able to train in high intensity (Rogers et al. 2021a; Rogers et al 2021b). However, it is still unknown whether the index of 0.75 remains stable in longer, constant load exercise intervals or in overtrained athletes (Rogers et al. 2021a) and if it still serves as a marker

for aerobic zone transition point (Rogers et al. 2021c). Also there remains questions whether the threshold is possible to be determined in field conditions, e.g., with chest belt or with different device's sample rates (Rogers et al. 2021a).

5.2 Assessment of second heart rate variability threshold

Cardiac control mechanisms of autonomic nervous system, which are activated when exercise intensity augments from moderate to high, aim for protecting cardiorespiratory system from harms like cardiac arterial desaturation or myocardial ischemia (Blasco-Lafarga et al. 2017). That activation can be seen in further reduced HRV (Cottin et al. 2006; Karapetian et al. 2008; Rogers et al. 2021b; Tulppo et al. 1996). It has been suggested that HRVT2 reflects suddenly increased mechanical and stretching influences on the heart. Furthermore, other mechanisms occurring in high exercise intensities, i.e., decreased diastolic filling time or diminished efficacy of Franck-Starling mechanisms might be involved in the appearance of HRVT2. Though, clear mechanisms behind the changes seen in HRV at that effort level are unclear. (Buccheit et al. 2007) In high exercise intensities, the reduction in HRV may be quick, and the resolution limits of traditional time and frequency domain methods may make it difficult to interpret (Sandercock & Brodie 2006). Hence, only one previous study where RMSSD would have been used for threshold determination were found. However, a few frequency domain and non-linear methods exist which are previously used for estimating second ET. Some of them are presented next.

Frequency domain analysis. In some studies, HF-power has been used for HRVT2 determination, as well. The curves of HF and HF*fHF measured during GXT have been used for HRVT2 determination, similarly as HRVT1. In those studies, HF*fHF values were either averaged every 20 s (Buccheit et al. 2007; Cottin et al. 2005; Cottin et al. 2006) or assessed using sliding 3 s time window (Mourot et al. 2014) during GXT and plotted against time and work intensity. The last abrupt increase in that curve has been determined as HRVT2. (Buccheit et al. 2007; Cottin et al. 2005; Cottin et al. 2006; Mourot et al. 2014). As with HRVT1, fHF alone seems not sufficient for threshold determination as clear nonlinear turning points are often impossible to be determined (Cottin et al. 2005; Cottin et al. 2006), maybe at least partly because of the effect of locomotor coupling (Di Michele et al. 2012).

There is evidence that short-term variability in R-R-interval lengths is more related to breathing patterns occurring during exercise than ANS activity status per se. Thus, both mean peak of HF power and absolute HF power density are affected by breathing frequency and tidal volume. (Buccheit et al. 2007) Cardiolocomotor coupling is another effect that may influence HRV power in when exercise intensity is augmented. The effect of cardiolocomotor coupling on HRV stays at the baseline level until the effort level of 60%max. After that, the effect has been shown to be the more pronounced the higher exercise intensity is at least in cycling. It is suggested that pedaling frequency component of HRV may account for 40 % of HRV when RSA accounts for the resting 60 %. (Blain et al. 2009) Hence, in some studies, the effect of respiratory sinus arrhythmia and the effect of locomotion have been taken into account by dividing HF power within exercise into two parameters; to one which takes into account the whole HF power spectrum (HF, 0.04-2 Hz) and to another, HF_{RSA}, which corresponds to frequencies ranging from 0.04 Hz to a cut-off frequency which makes the borderline between respiratory related and locomotion related components of HF. The aim for that division is to minimize the effect of mechanical respiratory-movement coupling on HRVT2 determination. The values of those both HF parameters, averaged from the last 60–90 s of each stage of GXT, have been plotted against workload and the onset point of the last increase in those curves have been determined as HRVT2. (Di Michele et al. 2012; Mendia-Iztueta et al. 2016; Mourot et al. 2014)

Poincaré plot. HRVT2 has been determined by visual inspection of the first breakpoint in SD2 curve, similarly as SD1 has been used for HRVT1 determination (Nascimento et al. 2017). Additionally, in one study, both SD1 and SD2 were tested for MLSS detection. They detected the minimal value of SD1 and the onset of plateaued phase by using 2 min moving averages from the end of GXT stages in overweight and obese people. (Schmidt et al. 2019) In another study both SD1 and RMSSD parameters were used for HRVT2-determination and the threshold was defined to GXT intensity in which the difference in successive values were first < 1 ms^2 (Novelli et al. 2018). During incremental exercise SD2 decreases progressively until the end and it is not only affected by PNS activation. Also, SD1/SD2-ratio have been seen to increase gradually during GXT and the most rapid increase begins from 80 %VO_2max, near second ET state. (Tulppo et al. 1996)

Non-linear analysis. DFA-a1 shows wider dynamic range than traditional parameters also in high intensities which makes it potential method for HRVT2 determination (Hautala et al. 2003; Mateo-March et al. 2022; Rogers et al. 2021b). When exercise intensity is further augmented

after first threshold intensity, progressive segregation of subsystems and mechanization of the whole system takes place (Gronwald et al. 2006). At high exercise intensities and under vagal blocking DFA-a1 values are strongly and linearly reduced (down to 0.3) (Hautala et al. 2003). That behavior reflects intensity-dependent changes in HR dynamics from strongly correlated to anticorrelated that are occurring because of vagal withdrawal and increase in sympathetic tonus. Also, other factors, e.g., coupling mechanisms of different physiological subsystems may have an influence on DFA-a1. (Gronwald et al. 2020)

DFA-a1 curve shows a slight trend change at about 80%max intensity, which is marked by blood lactate values, RER and RPE usually as second ET. (Blaco-Lafarga et al. 2022) HRVT2 has been determined in previous studies as a point in which DFA-a1 reach value of 0.5 (Rogers et al. 2021b; Mateo-March et al. 2022). That value of DFA-a1 have been chosen as it is suggested to reflect an intensity boundary just before maximal intensity zone, as does second ET defined from blood lactate or ventilatory parameters (Gronwald & Hoos 2019). ANS is not affecting HRV anymore at maximal intensities. Reduction in correlation properties of DFA-a1 (< 0.3) might reflect increased activation of complementary neural mechanisms and increased locomotor-respiratory coupling. Mechanical and neurological coupling effects on the sinus node are caused by movements of ribcage, blood pumping from leg muscles and pressure-wave changes and can be seen in values of non-linear parameters. (Blasco-Lafarga et al. 2017)

It is still unclear whether factors such as athlete's fitness level, study protocol, exercise type, food and caffeine intake and previous training would influence on HRVT2 determination. Some individual disparities have been found. For instance, the decrease in DFA-a1 values have been seen to be more pronounced when exercising at intensities over 70 %VO_2peak in medium and high-level trained individuals (Gronwald & Hoos 2019). Further validation is thus needed considering if the method is applicable for consumer devices. Also, limits of "tolerated" amounts of artifacts should be defined for all HRV methods. In addition, further studies especially in female population are needed. (Rogers et al. 2021b)

6 ASSOCIATION BETWEEN HRVTS AND TRADITIONAL THRESHOLDS

As already described, lactate concentration and ventilatory parameters have widely been used as "gold standard" methods for threshold assessment. Because of noninvasive and easy to implement nature of HRV threshold methods, they have been tested for alternatives to LTs and VTs. In many studies, first exercise threshold (ET) assessed from HRV data (HRVT1) has been found to correlate with VT1 (Mendia-Iztueta et al. 2016; Rogers et al. 2022; Sales et al. 2011) and LT1 (Mateo-March et al. 2022; Nascimento et al. 2017; Sales et al. 2011) while study populations have been quite small and, for instance, studies conducted in females are scarce.

HRVT1 determination using traditional time domain parameters and Poincaré plot -method has been successful in some studies. HRVT determined from cycling GXT using visual examination of deflection point or limit value of < 3 ms^2 from RMSSD and SD1 curves has been indicated to correlate with LT and VT in healthy adults. (Karapetian et al. 2008; Novelli et al. 2018; Queiroz et al. 2017) However, for example, in skiing HRVT1 determination from RMSSD and SD1 deflection points was successful only when considerable upper body movement not existed (Mendia-Iztueta et al. 2016). When HRVT1 was determined for young healthy males by using SD1 of Poincaré plot and RMSSD with the same criteria (deflection point or stabilization < 3 ms^2) Poincaré plot seemed to be more accurate method (Queiroz et al. 2017). HRVT1 assessed by using Poincaré plot analysis during maximal running test in recreational runners (Nascimento et al. 2017) and during incremental cycling test in diabetic and non-diabetic middle-aged people also showed good agreement with LT1 (determined using Dmax -method) and VT1 (Sales et al. 2011). However, in a study conducted in runners the absolute difference in running speed was 0.7 km/h on average which may be prominent in practice. (Nascimento et al. 2017)

VT1 has been successfully determined from the mathematical product of HF and fHF in running (Cottin et al. 2006) and cycling test (Cottin et al. 2005; Quinart et al. 2014). Instead, the method where only fHF parameter is used have been found unapplicable for running test (Cottin et al. 2006). In cycling the HF*fHF is also more recommendable for threshold determination than fHF only (Cottin et al. 2005). The effect of stride frequency on breathing increases at the same time with increases in running speed which results in linear relationship between fHF and running speed. As fHF and running speed are linearly related during incremental test, the detection of nonlinear increases in fHF or BF itself is impossible, even though it has been

suggested to be possible in some previous studies. Yet, by using the mathematical product HF*fHF the effect of tidal volume and vagal withdrawal (in addition to BF) can be detected. Additionally, when HF is multiplied by fHF the increase in both is amplified, thus, linear increase at threshold state is more easily detectable. (Cottin et al. 2006)

Strong linear correlations with VT1 and HRVT1 have been found in studies conducted for different study populations when DFA-a1 value 0.75 have been used as threshold marker (Rogers et al. 2021a; Rogers et al. 2021c). Either minimal differences ($p < 0.001$) between VT1 and HRVT1 presented as VO_2 or HR were detected in running test conducted for healthy people (Rogers et al. 2021a) or in cycling test conducted for cardiac disease population (Rogers et al. 2021c). Strong correlation between LT1 (assessed using log-log method) and HRVT1 for both HR and power were also found when elite triathletes (Rogers et al. 2022) and elite endurance cyclists were tested (Mateo-March et al. 2022). Subjects of these studies differ by age, weight, and fitness level. Thus, the results are representing general population quite well. It is still worth mentioning that in those aforementioned studies, there were either no female participants at all (Mateo-March et al. 2022; Rogers et al. 2021a; Rogers et al. 2021c) or female subjects were minority (Rogers et al. 2022).

HF, RMSSD and SD1 have been used for first ET determination by observing the point when they reach a nadir. However, the determination of second threshold from those parameters is problematic because there seldom exists any further dynamic range after the first threshold. (Cottin et al. 2006; Karapetian et al. 2008; Tulppo et al. 1996) For example, in overweight and obese people HRVT detected using either SD1 or SD2 was overestimating MLSS (Scmidt et al. 2019). In addition to that, while traditional HRV parameters decline when exercise intensity increases, in many individuals a detectable nadir is not reached before maximal intensities (Rogers et al. 2021c). Even if HRVT2 can be detected the reproducibility might be weaker when compared to HRVT1 (Novelli et al. 2018). Also, an increase in artefact levels occur relative to exercise intensity when intensity is augmented to 80–100 %VO_2max, maybe at least partly because of increased HR and chest movements (Giles & Draper 2018). Especially exercise modes where upper body is used, significant error is caused by the effects of locomotor to the breathing frequency and HF band. Thus, methods in which HRVT2 is determined using fHF only have been noticed to be unsuitable for sports with upper body involved (e.g., skiing and swimming) (Mendia-Iztueta et al. 2016; Mourot et al. 2014).

Even if those aforementioned artifact sources with increasing exercise intensity are evident, in some studies HRVT2 determination from traditional parameters have been successful. When running speed at HRVT2 have been compared with running speed at VT2 no significant differences have been found in male soccer players (Cottin et al. 2006) or in trained boys (Buccheit et al. 2007). Significant differences were neither noticed in VO_2 or HR at HRVT2 when compared to VT2 when male triathletes/competitive cyclists (Cottin et al. 2005) or obese adolescents (Quinart et al. 2014) were tested. In those above-mentioned studies, HRVT2 has been determined as the test timepoint at which second and last nonlinear increase in HF*fHF mathematical product appear (Buccheit et al. 2007; Cottin et al. 2005; Cottin et al. 2005; Quinart et al. 2014). That nonlinear increase at about the same time at which VT2 occur is induced by the mechanical effect of the increase in BF on the sinus node combined with increased tidal volume (Cottin et al. 2006). HRVT2 has also been successfully determined using slightly different mathematical HF-derived product where total HF power spectrum is divided into respiratory and locomotor components (Mendia-Izueta et al. 2016; Mourot et al. 2014). With that mathematical calculation, no significant differences in speed or HR at HRVT2 compared to VT2 were noticed in skiing (Mourot et al. 2014) or Nordic walking (Mendia-Izueta et al. 2016). Neither significant differences were noticed in HR or speed in HRVT2 when compared to LT2 in swimmers (Di Michele et al. 2012). The calculation of locomotor component of HF seems to further increase the accuracy of HRVT2 to estimate VT2 at least in sports where upper body is involved (Mourot et al. 2014). With respect to locomotion, the test mode affects HRVT2 determination as, for example, in cycling pedaling rate is usually kept constant during GXT stages whereas in running stride frequency increases with intensity increments affecting on BF (Buccheit et al. 2007).

When HRVT2 has been determined using DFA-a1, power output and HR at LT2 turned out to be significantly higher than at HRVT2 (DFA-a1 0.5). In that study LT2 was determined as a workload when there appeared > 2 mmol/l rise in lactate compared to baseline. The subjects were elite endurance cyclists. Thus, positive correlations were found between power output and HR at LT2 and HRVT2 which might reflect that both measurements are showing a constant error. (Mateo-March et al. 2022) Rogers et al. (2021b) found out that HR at that point when DFA-a1 value reached 0.5 in running GXT was closely related to HR at VT2 in recreational male runners with wide age and fitness level (Rogers et al. 2021b). Further studies regarding to the comparison of HRVT2 and LT2 with different methods are still needed. (Mateo-March et al. 2022)

7 RESEARCH QUESTIONS

The purpose of this study was to test the applicability of HRVT methods presented in previous studies in nonathlete female runners, and in running test.

Question 1: Is it possible to determine first and second exercise threshold (ET) from heart rate variability data measured during maximal, incremental running test in novice female runners?

Hypothesis: In most cases, yes. In previous studies HRVT1 determination has been successful and exclusion rates low using traditional HRV-parameters, RMSSD and SD1 (Cambri et al. 2016; Karapetian et al. 2008; Mendia-Izueta et al. 2016; Novelli et al. 2018; Queiroz et al. 2017; Sales et al. 2011) although body movements, that might be occurring in running also, may cause errors to data (Mendia-Iztueta et al. 2016). HRVT1 determination from the mathematical product of HF and fHF (HF*fHF) has been mostly successful in running when athletes and obese adolescents have been tested and the threshold has been unidentifiable for only few participants (Cottin et al. 2006; Quinart et al. 2014). Thus, it is possibly suitable method for novice runners in this current study, too. When HRVT1 have been determined using the turning point of DFA-a1 value 0,75 in elite cyclists (Mateo-March 2022), in individuals with cardiac disease (Rogers et al. 2021c) and in healthy adult males (Rogers et al. 2021a) the threshold have been identifiable for all the participants. It is possible that the threshold is then similarly detectable for healthy female runners. Yet, the sample duration was shorter in the present study than in previous studies which may affect the results.

Determination of second threshold from traditional HRV parameters (RMSSD, HF) may be problematic or even impossible as there seldom exists any further dynamic range when the intensity is increased above first threshold. (Cottin et al. 2006; Karapetian et al. 2008; Tulppo et al. 1996) In addition, it has been suggested that for many individuals a detectable flattening of the values is not reached before maximal intensities which may result in difficulties in detecting both thresholds (Rogers et al. 2021c). When the effect of respiratory sinus arrhythmia on HF-variability is taken into account the HRVT2 might be detectable from HF-variability (Di Michele et al. 2012). Determination of HRVT2 using DFA-a1 value 0,5 seems applicable method at least in athlete populations (Mateo-March et al. 2022; Rogers et al. 2021b) but it is unclear is the same method suitable for non-athletes, as well.

Question 2: Is it possible to accurately estimate LT1 and LT2 (as HR and running speed) by different HRVT1 and HRVT2 methods, respectively?

Hypothesis: It may vary depending on the method used. VO_2 at HRVT1 determined using RMSSD and SD1 curves has been indicated to correlate with VO_2 at LT1 in healthy adults. (Karapetian et al. 2008) Exercise intensity at HRVT1 determined using Poincaré plot analysis in recreational runners (Nascimento et al. 2017) and in middle-aged healthy and diabetic people have also shown good agreement with LT1 (Sales et al. 2011). Strong correlations between heart rate, running speed and cycling power at HRVT1 determined using mathematical product of HF-power and VT1 in different study populations has also been found (Cottin et al. 2006; Quinart et al. 2014). Further, when DFA-a1 threshold of 0,75 has been used, minimal differences between VT1 and HRVT1 when expressed as VO_2 and HR were detected either in running test conducted for healthy, non-athletic population (Rogers et al. 2021a) or cycling test conducted for cardiac disease population (Rogers et al. 2021c). No differences between exercise intensity and HR at LT1 and HRVT1 have been found when elite triathletes have been tested (Mateo-March et al. 2022; Rogers et al. 2022). Thus, it is presumable that correlations would be found in novice runners, too, although there are hardly any previous research data about females.

Previous evidence about the correlations between HRVT2 and LT2 is more controversial. Running speed at VT2 and HRVT2 has not been significantly different when HRVT2 has been determined from HF*fHF curve in athlete populations (Buccheit et al. 2007; Cottin et al. 2006). Neither significant difference in VO_2 or HR have been noticed when same method have been used and HRVT2 compared to VT2 in cycling (Cottin et al. 2005). However, when movements in upper body are occurring that method may overestimate HR at threshold (Mourot et al. 2014). Power output and HR at HRVT2 determined to the point at which DFA-a1 reach value 0,5 have been shown to be significantly lower than at LT2 (Mateo-March et al. 2022). In recreational male runners of different ages and fitness levels, HR at DFA-a1 derived HRVT2 was instead closely related to HR at VT2 (Rogers et al. 2021b). Thus, it would be suggested that in recreational female runners the relationship would be similar.

Question 3: Are different HRVT methods applicable in running GXT protocol conducted according to the Finnish guidelines?

Hypothesis: They may be but e.g., treadmill stopping for lactate sampling between stages may cause bias. HRV during especially high intensity exercise is prone to bias thus the pauses may have affected physiological functions and can be seen in HRV (Sandercock & Brodie 2006). Previous studies have shown that HRV recovery is faster the lower exercise intensity is and the fitter subject is (Stanley et al. 2013). Thus, considering this, the 10 s pause would have not maybe been long enough to cause detectable changes in HRV at least in higher intensities in this study population. However, in many previous studies where running GXT protocol have been used it has been conducted in a continuous manner and treadmill has not been stopped between stages (Buccheit et al. 2007; Cottin et al. 2006; Nascimento et al. 2017; Rogers et al. 2021a & 2021b). In previous studies also the stage duration of running test has been shorter and 1 min stages have been used (Buccheit et al. 2007; Cottin et al. 2006). Although also 3 min (Nascimento et al. 2017; Rogers et al. 2021a & 2021b) and even 4 min (Blasco-Lafarga et al. 2017) stages have been used in some research protocols. The variation of stage duration makes the comparisons more difficult.

8 METHODS

In this section the measurements and protocols used are presented. All measurements for this study were conducted in Tampere Research Center of Sports Medicine in UKK Institute (Tampere, Finland) during March-April 2021. The study protocol and methods were accepted by the ethics committee of Pirkanmaa Hospital District.

8.1 Participants

Study subject group consisted of 24 female participants. They were recruited from the participants of a larger running related RCT-study conducted by Tampere Research Center of Sports Medicine and UKK Institute. All participants volunteered for the study. The study started with initiation measurements in March 2021. Altogether, tests were conducted to 28 participants but data from 4 participants was excluded due to high number of artefacts in HRV signal. Participants were selected according to certain criteria: 20–40 years old females with no chronic diseases or regular medications. All the participants were novice runners with running as a main sport. They had for maximum two years' experience of regular running with less than 15 km/week.

The characteristics of the study participants are presented in table 1. The percentage of body fat from total body mass was estimated using skinfold thickness measurement according to four-site method by Durnin and Womersley (1974). VO_2max was measured in treadmill GXT.

TABLE 1. The characteristics of the participants (n = 24).

	Age (y)	Height (cm)	Body mass (kg)	BMI	Fat-%	VO_2max (ml/kg/min)
average	33,3	166,7	67,0	24,1	28,9	38,9
SD	4,5	7,1	12,7	4,4	4,4	4,2

SD = standard deviation, BMI = body mass index

8.2 Test protocol and methods

Test protocol consisted of resting measurements before exercise, maximal graded exercise test in treadmill and recovery measurements. All the tests were conducted in the morning between 9–11 a.m. Participants were informed about the measurement protocol beforehand. They were also instructed to avoid heavy training for two days, alcohol for 24 hours and caffeine for 2 hours before the test. For the reliability of GXT results, the circumstances of the preceding day should be standardized, and it is recommended to avoid heavy and high-volume exercise two days before test day (Nummela & Peltonen 2018, 80). Also, if they had had any respiratory infection, they should have had 7 days without any symptoms before attending to the test. R-R-interval lengths and heart rate were measured continuously during the whole test protocol by using Firstbeat Bodyguard 2 -monitor (Firstbeat Technologies Ltd., Jyväskylä, Finland). Heart functions were also monitored with 12-lead ECG-device (Case 6.5, GE Healthcare, Milwaukee Wisconsin, USA). ECG-electrode placement was conducted according to international guidelines for 12-lead ECG-electrode placement. Limb electrodes were positioned on torso near proximal ends of the limb for their placing to kept as still as possible during running. Blood lactate concentration was measured using capillary blood samples taken from fingertip at rest and at the end of each stage. The samples were analyzed using Biosen C–line Clinic -analyzer (EKF Diagnostics, Barleben, Germany). Before the start of the test subjects were instructed to tell or show with hand signs as soon as they want to stop the test.

The preparations and basic measurements were conducted according to the Finnish exercise testing guide (Nummela & Peltonen 2018, 80). Test room conditions, temperature (°C), air pressure (mmHg) and air humidity (%) were recorded. Before the test participant was asked to read and sign in a written consent form. After that anthropometric measurements, height (cm), body mass (kg) and body composition (skinfold caliper) were measured. Also resting ECG and blood pressure were measured at rest before the initiation of the actual measurement protocol. Firstbeat Bodyguard 2 -monitor was placed on chest and RR-interval monitoring was started. Resting measurements were conducted in laying position for 10 minutes and after that in sitting position for 5 minutes. Participant was instructed to sit without speaking or moving. After 15 min resting measurements ECG-transmitter was placed on subject’s hips with belt, subject was instructed to step on a treadmill (Telineyhtymä, Kotka, Finland) and treadmill safety harness was fastened. The first blood sample was taken for resting lactate determination. Ventilatory

parameters were measured throughout the test via breath-by-breath gas analyzer (Vyntus CPX, Vyaire Medical GmbH, Höchberg, Germany).

Before the initial GXT protocol, a 3 min warm up was performed with treadmill speed set to 5,0 km/h and inclination 0,6 °. Immediately after warming up the first stage of GXT with the speed of 5,0 or 6,0 km/h was started. Starting speed was individually chosen by the testing personnel based on participant's running experience which they have reported in the background information form (appendix 1). Running speed was increased progressively by 1 km/h every 3 min. This kind of continuous testing protocol with 3 min stages has been previously proposed to be valid and reliable method for LT determination (Weltman et al. 1990). HR was determined as the average of last 15 s for each 3 min stage from ECG-monitor. After each 3 min stage the treadmill was stopped for blood sampling. Samples were put into lactate analyzer which measured the lactate concentration of each individual sample. The test was continued until subjective exhaustion. Immediately after that, the last blood sample was taken, and the subject was detached from the respiratory gas analyzer and safety harness.

8.3 Data analysis

R–R-interval data measured during GXT was analyzed with Firstbeat Sports v.4.7. –program (Firstbeat Technologies Ltd., Jyväskylä, Finland) and Kubios HRV Standard 3.5.0 (Kubios Oy, Kuopio, Finland). For frequency domain analysis Fast Fourier transform -method was used with window width of 60 s and window overlap of 50 %. HF–band width was set to 0,15-1 Hz. DFA-a1 was analyzed using detrended fluctuation analysis. All Kubios analysis were conducted with default settings. Detrending method was kept in Smoothn priors (lambda: 500) and interpolation rate as 4 Hz.

The values of all HRV-parameters were averaged for the last 60 s of each stage as it has been suggested to be long enough for accurate determination of RMSSD (Goldberger et al. 2006) and HF-parameters (Task Force 1996). One minute recording phases have also previously been used for Poincaré plot -analysis (Nascimento et al. 2019). DFA-a1 window width was kept at 4-16 beats as suggested by previous studies (Gronwald et al. 2020). For DFA-a1 previous evidence recommends 2 min sliding analysis windows with 50 % overlap (Hardstone et al. 2012) but for more accurate comparison of the results obtained with different methods the

analysis windows of the same length were decided to use. It was also of interest if 1 min averaged values of DFA-a1 are applicable for that kind of settings. The recording phases for analysis were taken from the end of the stage for obtaining as stationary data as possible. If notable amount of bias and/or false beats were seen in the last 60 s of the stage, the analysis window length was either shortened from either end or the window was moved to a slightly earlier moment of that stage. These minor change in some individuals would possibly have had no effect on the results as recording timing inside the stage has been seen to have little effect on values (Hautala et al. 2003). Respiratory parameters were measured breath by breath and averaged for the last 30 s of each stage.

Kubios artefact correction level was kept at 3 % and at medium level. Previous studies suggest the use of correction level 1-3 %. Even though it may cause bias into DFA-a1 values in high exercise intensities there is evidence that it has no effect on HRVT determination (Rogers et al. 2021d).

8.4 Lactate threshold determination

Lactate thresholds were determined according to the Finnish guidelines (figure 5, Nummela & Peltonen 2018, 97). LT1 was determined to the lactate curve of the GXT to the point at which lactate concentration reached value which was +0,3 mmol/l of its lowest value measured during the test. The accuracy of half a test stage was used. It was then plotted against speed and HR. For LT2 determination two linear lines were fitted to the lactate curve. One was set to intersect the point of the curve in which LT1 occurs and the lactate value of the next stage. The other linear fitting was set to traverse the lactate points of the last test stages among which the increase in lactate concentration was above 0,8 mmol/l. The test stage or the halfway mark of two stages which were located nearest to the intersection of those two lines were then determined as LT2. LT2 was similarly reported as speed and HR as those are reported to be practical and reliable threshold measures (Pfitzinger and Freedson 1998).

Lactate thresholds were identified by two investigators. One of them first determined the thresholds which the other, more experienced investigator checked afterwards. If investigators disagreed to some LT also ventilatory parameters were checked and the threshold was set as accurately as possible by taking account, both changes in lactate concentration and ventilation.

8.5 Heart rate variability threshold determination

Heart rate variability thresholds were determined using four different analysis methods: time domain analysis (RMSSD), frequency domain analysis (HF*fHF), Poincaré plot (SD1 and SD2) and detrended fluctuation analysis (DFA-a1). All thresholds were determined as speed and heart rate with an accuracy of half of a test stage. If the value of HRV-parameter reached the threshold in between two test stages, the threshold speed and HR were determined as the midpoint of the values of those two stages.

$HRVT1_{RMSSD}$ was determined using the values of RMSSD-parameter of time domain analysis measured during the treadmill test. $HRVT1_{RMSSD}$ was set to the test stage or to the halfway of two stages in which the values dropped below 3,0 ms^2 (Karapetian et al. 2008; Mendia-Izueta et al. 2016). Heart rate and speed at that point were then determined as threshold. The determination of $HRVT1_{RMSSD}$ is described in figure 7.

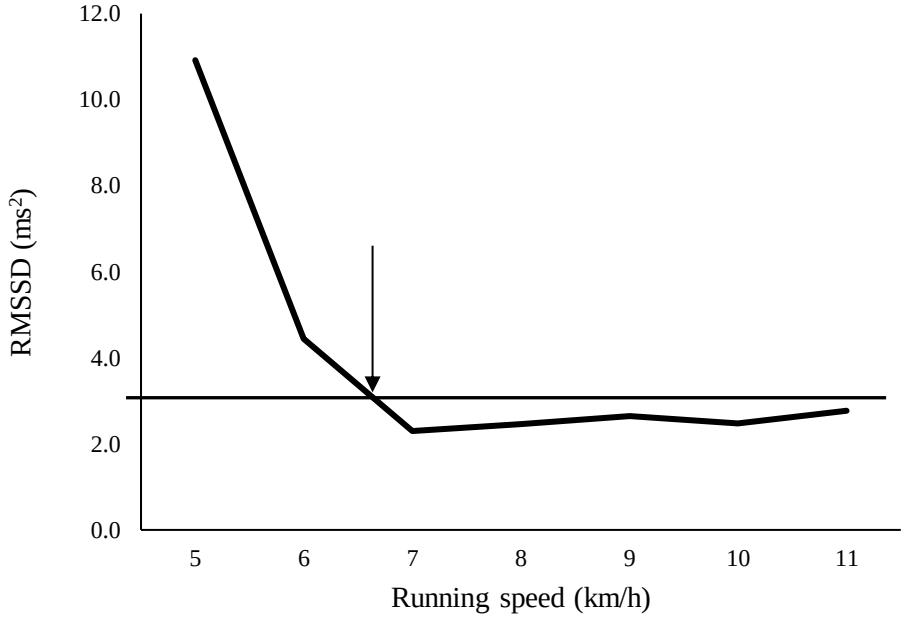


FIGURE 7. The determination of HRVT1 of one subject from the values of RMSSD-parameter measured during GXT.

$HRVT1_{HF}$ was determined by using mathematical product of HF*fHF as suggested by previous studies (Cottin et al. 2005; Cottin et al. 2006). Mathematical product was calculated by multiplying absolute HF-power (ms^2) by HF peak frequency (Hz). The values were further

modified into natural logarithm by using function y = ln (1 + x) in Microsoft Excel (Microsoft Corporation, Redmond, USA) for obtaining data with normal distribution. $HRVT1_{HF}$ was determined to the timepoint of the test where first nonlinear increase in HF*fHF curve occurred (Cottin et al. 2005; Cottin et al. 2006; Quinart et al. 2014). $HRVT2_{HF}$ was determined by using that same curve to a test stage (or halfway of stages) when the last abrupt increase occurred (Buccheit et al. 2007; Cottin et al. 2005; Cottin et al. 2006; Mourot et al. 2014). The determination of $HRVT1_{HF}$ and $HRVT2_{HF}$ from HF*fHF-curve is depicted in figure 8.

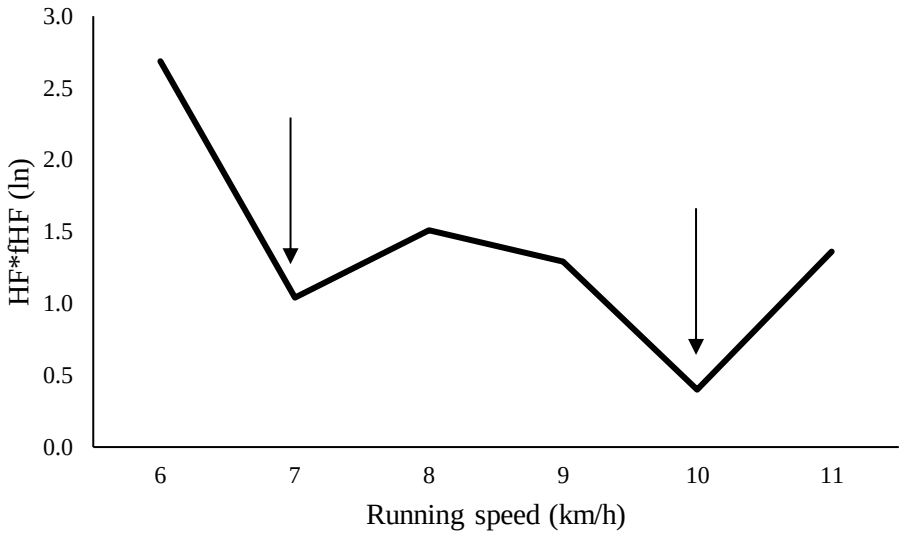


FIGURE 8. The determination of HRVT1 and HRVT2 of one subject from the values of HF*fHF-mathematical product measured during GXT.

Poincaré plot -analysis and its parameters, SD1 and SD2 were used for $HRVT1_{SD1}$ and $HRVT2_{SD2}$ determination, respectively. $HRVT1_{SD1}$ was determined to the point of the SD1-curve at which SD1 reached lowest values and after which the values were plateaued below 3,0 ms^2 during GXT (figure 9). The method was same that has been used in previous studies (Sales et al. 2011). Correspondingly, $HRVT2_{SD2}$ was determined from SD2-curve following the same criteria (figure 10).

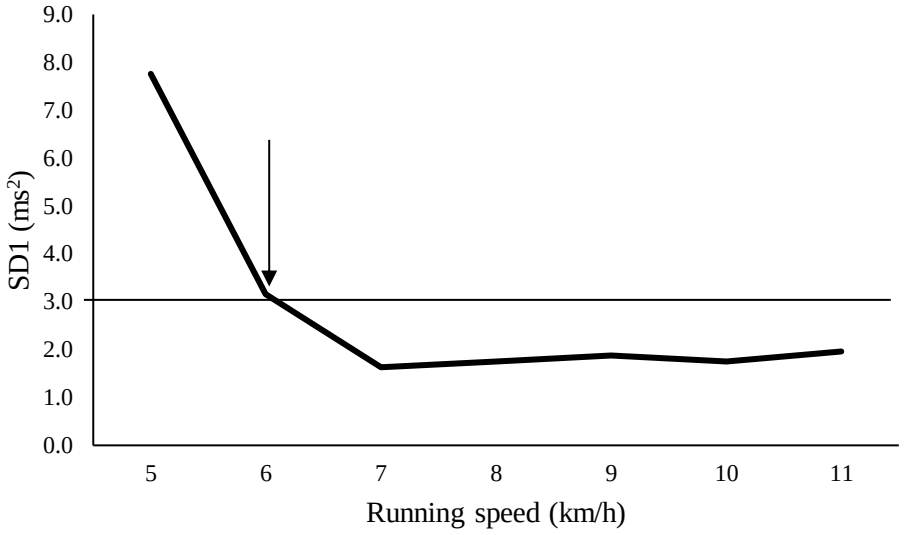


FIGURE 9. The determination of HRVT1 from the values of SD1 parameter of one subject measured during GXT.

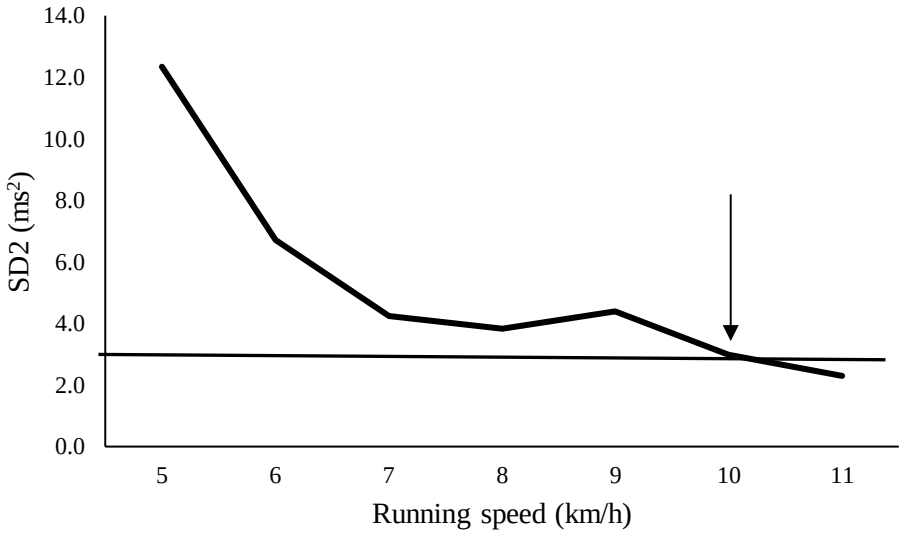


FIGURE 10. The determination of HRVT2 from the values of SD2 parameter of one subject measured during GXT.

$HRVT1_{DFA}$ was determined as the point in which DFA-a1 values measured during the test first reached value of 0,75 as suggested previously by a couple of studies (Rogers et al. 2021a; Rogers et al. 2022; Mateo-March et al. 2022). Subsequently, $HRVT2_{DFA}$ was set to a test stage in which the values were first lowered below 0,5 (Rogers et al. 2021b; Mateo-March et al. 2022). The determination is presented in figure 11. DFA-a1 derived thresholds were also determined with an accuracy of a half test stage.

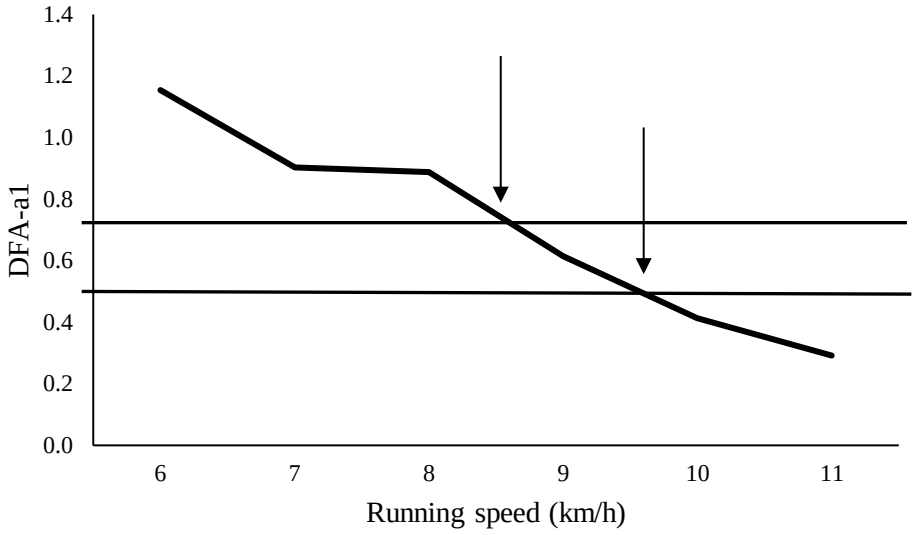


FIGURE 11. The determination of HRVT1 and HRVT2 of one subject from the values of DFA-a1-parameter measured during GXT.

8.6 Statistical analysis

Statistical analyses were conducted using IBM SPSS Statistics 26 -program (International Business Machines Corporation, Armonk, USA). For additional analysis, such as, linear regression charts and tables also Microsoft Excel for Mac 16.49 -application (Microsoft Corporation, Redmond, USA) was used.

Normality of the analyzed parameters was tested with Shapiro-Wilk test. Correlations within HRVTs and LTs were calculated with Spearman's rho separately for each HRVT method used. Linear regression graphs were created by plotting HR or speed at LT as a function of the value of the same parameter at HRVT. Strength of linear relationships was estimated from correlation coefficient values as follows; very strong when $r \geq 0,8$, moderately strong when $0,6 \leq r < 0,8$, fair when $0,3 \leq r < 0,5$ and low when $r < 0,3$ (Chan 2003). The statistical significance was accepted as $p \leq 0,05$ in all tests.

Whereas the values of all parameters were not shown to be normally distributed according to Shapiro Wilk's test, non-parametric Related-Samples Wilcoxon Signed Rank Test was used to test paired differences of HR and speed values. Bland-Altman plot charts were created using

IBM SPSS Statistics 26 -program (International Business Machines Corporation, Armonk, USA) with 95 % limits of agreement (LOA) according to Bland and Altman (1999).

Correlations between thresholds assessed using different methods were also examined using intraclass correlation coefficients (ICCs). ICC estimates with 95 % confident intervals were set to be based on absolute agreement of single measurement and 2-way mixed-effects model as suggested by Koo and Li (2016). ICC values were interpreted according to the same study and values of less than 0,5 were considered to reflect poor, values between 0,5–0,75 moderate, values between 0,75–0,9 good and values above 0,9 excellent reliability. Values are reported as average ± SD. (Koo and Li 2016)

9 RESULTS

In this section, the results of the study are represented. LT1 and LT2 were successfully determined for all participants. $HRVT1_{SD1}$, $HRVT2_{SD2}$, $HRVT1_{HF}$ and $HRVT1_{DFA}$ were detectable for all 24 participants. $HRVT1_{RMSSD}$ was detectable for 23 participants, $HRVT2_{DFA}$ for 20 participants and $HRVT2_{HF}$ for 14 participants.

9.1 Lactate and HRV thresholds

Thresholds were reported as speed and heart rate. The participants' average heart rate at LT1 was 144,3 ± 2,2 bpm, speed 6,8 ± 0,2 km/h and lactate concentration 1,4 ± 0,1 mmol/l. At LT2 heart rate reached 173,5 ± 1,6 bpm, speed 9,0 ± 0,2 km/h and lactate concentration 3,3 ± 0,1 mmol/l. Means, and standard deviations of lactate thresholds of the subject group are presented in table 2.

TABLE 2. The means and standard deviations of lactate thresholds of the study participants determined as HR, speed and blood lactate concentration.

		Mean	SD
LT1	HR (bpm)	144,3	2,2
	Speed (km/h)	6,8	0,2
	Lactate concentration (mmol/l)	1,4	0,1
LT2	HR (bpm)	173,5	1,6
	Speed (km/h)	9,0	0,2
	Lactate concentration (mmol/l)	3,3	0,1

LT1 = first lactate threshold, LT2 = second lactate threshold, HR = heart rate, bpm = beats per minute, km/h = kilometers per hour, mmol/l = millimoles per liter, SD = standard deviation

HRV-thresholds were reported as heart rate and speed similarly as lactate thresholds. Heart rate at first exercise threshold (ET) ranged from 136 ± 2,6 to 165 ± 3,8 bpm depending on method used. Running speeds at HRVT1 varied from 6,4 ± 0,2 to 8,3 ± 0,3 km/h and lactate concentrations from 1,2 ± 0,5 to 2,9 ± 1,5 mmol/l. At HRVT2 heart rate of the participant group ranged between 162 ± 2,4 and 182 ± 3,6 bpm, speed between 7,9 ± 0,2 and 10,7 ± 0,4 km/h and lactate concentration between 2,4 ± 1,0 and 5,3 ± 1,8 mmol/l. Mean values and standard deviations of HRV thresholds are presented in table 3.

TABLE 3. The means and standard deviations of HRV thresholds of the study participants determined as HR, speed and blood lactate concentration.

		Mean	SD
$HRVT1_{RMSSD}$	HR (bpm)	146	3,3
	Speed (km/h)	6,8	0,2
	Lactate concentration (mmol/l)	1,6	0,7
$HRVT1_{HF}$	HR (bpm)	164	3,8
	Speed (km/h)	8,1	0,3
	Lactate concentration (mmol/l)	2,9	1,5
$HRVT1_{SD1}$	HR (bpm)	137	2,6
	Speed (km/h)	6,4	0,2
	Lactate concentration (mmol/l)	1,2	0,5
$HRVT1_{DFA}$	HR (bpm)	165	3,5
	Speed (km/h)	8,3	0,3
	Lactate concentration (mmol/l)	2,8	1,0
$HRVT2_{HF}$	HR (bpm)	182	2,4
	Speed (km/h)	10,7	0,2
	Lactate concentration (mmol/l)	5,3	1,7
$HRVT2_{SD2}$	HR (bpm)	162	2,4
	Speed (km/h)	7,9	0,2
	Lactate concentration (mmol/l)	2,4	1,0
$HRVT2_{DFA}$	HR (bpm)	177	3,6
	Speed (km/h)	9,6	0,4
	Lactate concentration (mmol/l)	4,8	1,8

$HRVT1_{RMSSD}$ = first heart rate variability threshold determined from RMSSD, $HRVT1_{HF}$ = first heart rate variability threshold determined from HF*fHF, $HRVT1_{SD1}$ = first heart rate variability threshold determined from SD1, $HRVT1_{DFA}$ = first heart rate variability threshold determined from DFA-a1, $HRVT2_{HF}$ = second heart rate variability threshold determined from HF*fHF, $HRVT2_{SD2}$ = second heart rate variability threshold determined from SD2, $HRVT2_{DFA}$ = second heart rate variability threshold determined from DFA-a1, HR = heart rate, bpm = beats per minute, km/h = kilometers per hour, mmol/l = millimoles per liter, SD = standard deviation

9.2 Comparison of lactate and HRV thresholds

Significant correlations ($p < 0,001$) were found in speed between LT1 and $HRVT1_{SD1}$ as well as between LT2 and $HRVT2_{SD2}$. Further less significant ($p < 0,05$) correlations were also found in HR between LT1 and $HRVT1_{HF}$, LT2 and $HRVT2_{HF}$ as well as between LT2 and $HRVT2_{DFA}$. Less significant correlations ($p < 0,05$) were also found when speed between LT2 and $HRVT2_{HF}$ was compared as well as between LT2 and $HRVT2_{DFA}$. Spearman's rho values of the LT–HRVT correlations are presented in table 4.

TABLE 4. Correlations between LT1 and HRVT1 and between LT2 and HRVT2. Thresholds are presented as HR and speed.

Spearman's rho	HR	speed
LT1 & $HRVT1_{RMSSD}$	0,337	0,286
LT1 & $HRVT1_{HF}$	0,436*	0,148
LT1 & $HRVT1_{SD1}$	0,352	0,633**
LT1 & $HRVT1_{DFA}$	0,089	0,348
LT2 & $HRVT2_{HF}$	0,618*	0,543*
LT2 & $HRVT2_{SD2}$	0,366	0,584**
LT2 & $HRVT2_{DFA}$	0,553*	0,449*

LT1 = first lactate threshold, LT2 = second lactate threshold, $HRVT1_{RMSSD}$ = first heart rate variability threshold determined from RMSSD, $HRVT1_{HF}$ = first heart rate variability threshold determined from HF*fHF, $HRVT1_{SD1}$ = first heart rate variability threshold determined from SD1, $HRVT1_{DFA}$ = first heart rate variability threshold determined from DFA-a1, $HRVT2_{HF}$ = second heart rate variability threshold determined from HF*fHF, $HRVT2_{SD2}$ = second heart rate variability threshold determined from SD2, $HRVT2_{DFA}$ = second heart rate variability threshold determined from DFA-a1, HR = heart rate

* $p < 0,05$, ** $p < 0,001$

Linear regression graphs can be found from appendix 2 (HR) and from appendix 3 (speed). They are showing the highest relationship between LT2 and $HRVT2_{HF}$ ($r = 0,68$) when considering HR and between LT1 and $HRVT1_{RMSSD}$ ($r = 0,48$) as well as between LT1 and $HRVT1_{SD1}$ ($r = 0,49$) regarding speed. No relationships between other pairs were found.

Differences between HR and speed values in LT and HRVT were also compared using Wilcoxon signed rank test. The analysis showed significant differences ($p < 0,001$) and high standard errors in HR between LT1 and $HRVT1_{HF}$, LT1 and $HRVT1_{DFA}$ and LT2 and $HRVT2_{SD2}$. Standard errors of those pairs were 28,8, 35,0 and 26,8 correspondingly. The p-value was apparently highest when LT1 and $HRVT1_{RMSSD}$ were compared (0,777), although standard error was still quite large (22,9). The results are presented in table 5 as regards to HR.

TABLE 5. Differences in HR at LTs and HRVTs determined by different methods according to Wilcoxon signed rank test.

Difference in HR between	W	Sig. (2-sided test)	SE
LT1 & $HRVT1_{RMSSD}$	0,283	0,777	22,9
LT1 & $HRVT1_{HF}$	3,530	< 0,001**	28,8
LT1 & $HRVT1_{SD1}$	-2,197	0,028*	24,8
LT1 & $HRVT1_{DFA}$	3,416	< 0,001**	35,0
LT2 & $HRVT2_{HF}$	3,063	0,002*	12,7
LT2 & $HRVT2_{SD2}$	-3,622	< 0,001**	26,8
LT2 & $HRVT2_{DFA}$	1,862	0,063	19,3

HR = heart rate, LT1 = first lactate threshold, LT2 = second lactate threshold, $HRVT1_{RMSSD}$ = first heart rate variability threshold determined from RMSSD, $HRVT1_{HF}$ = first heart rate variability threshold determined from HF*fHF, $HRVT1_{SD1}$ = first heart rate variability threshold determined from SD1, $HRVT1_{DFA}$ = first heart rate variability threshold determined from DFA-a1, $HRVT2_{HF}$ = second heart rate variability threshold determined from HF*fHF, $HRVT2_{SD2}$ = second heart rate variability threshold determined from SD2, $HRVT2_{DFA}$ = second heart rate variability threshold determined from DFA-a1, W = standardized test statistic of Wilcoxon signed rank test, SE = standard error

* $p < 0,05$, ** $p < 0,001$

The differences in speed between analyzed pairs according to Wilcoxon signed rank test are presented in table 6. The same three pairs, LT1 and $HRVT1_{HF}$, LT1 and $HRVT1_{DFA}$ and LT2 and $HRVT2_{SD2}$, also showed significant differences ($p < 0,001$) in speed as in HR. The standard errors of those pairs were also quite large (28,6, 34,9, 26,6, correspondingly). Additionally, significance of paired difference between LT1 and $HRVT1_{RMSSD}$ was again the most remarkable ($p = 0,926$, SE = 22,4).

TABLE 6. Differences in HR at LTs and HRVTs determined by different methods according to Wilcoxon signed rank test.

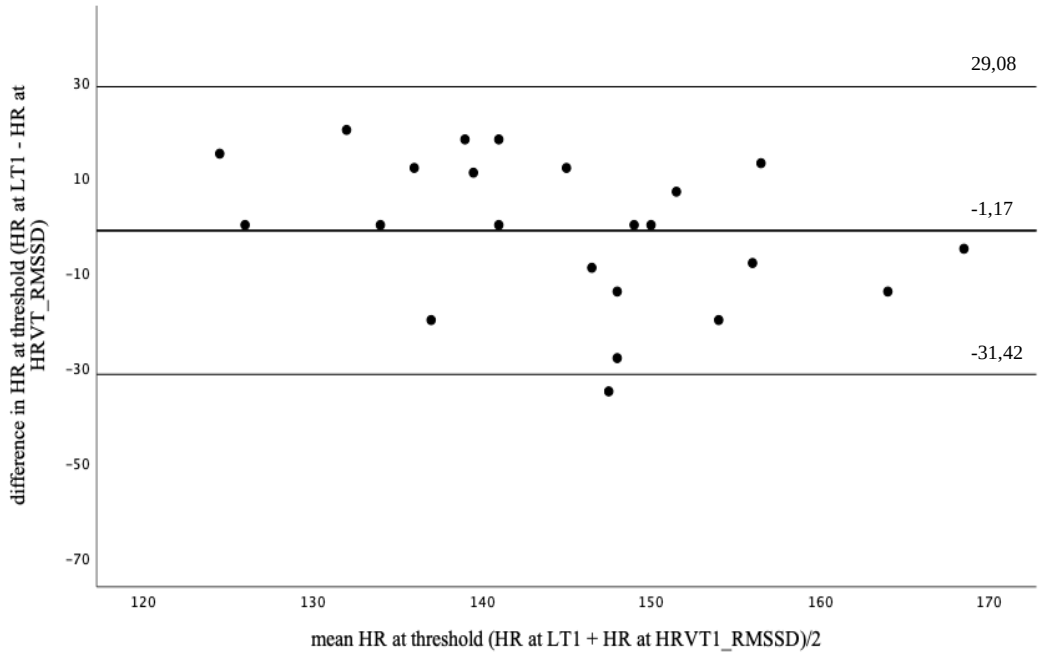
Difference in speed between	W	Sig. (2-sided test)	SE
LT1 & $HRVT1_{RMSSD}$	0,089	0,929	22,4
LT1 & $HRVT1_{HF}$	3,701	< 0,001**	28,6
LT1 & $HRVT1_{SD1}$	-2,384	0,017*	24,3
LT1 & $HRVT1_{DFA}$	3,543	< 0,001**	34,9
LT2 & $HRVT2_{HF}$	3,089	0,002*	12,6
LT2 & $HRVT2_{SD2}$	-3,568	< 0,001**	26,6
LT2 & $HRVT2_{DFA}$	1,921	0,055	19,3

LT1 = first lactate threshold, LT2 = second lactate threshold, $HRVT1_{RMSSD}$ = first heart rate variability threshold determined from RMSSD, $HRVT1_{HF}$ = first heart rate variability threshold determined from HF*fHF, $HRVT1_{SD1}$ = first heart rate variability threshold determined from SD1, $HRVT1_{DFA}$ = first heart rate variability threshold determined from DFA-a1, $HRVT2_{HF}$ = second heart rate variability threshold determined from HF*fHF, $HRVT2_{SD2}$ = second heart rate variability threshold determined from SD2, $HRVT2_{DFA}$ = second heart rate variability threshold determined from DFA-a1, W = standardized test statistic of Wilcoxon signed rank test, SE = standard error

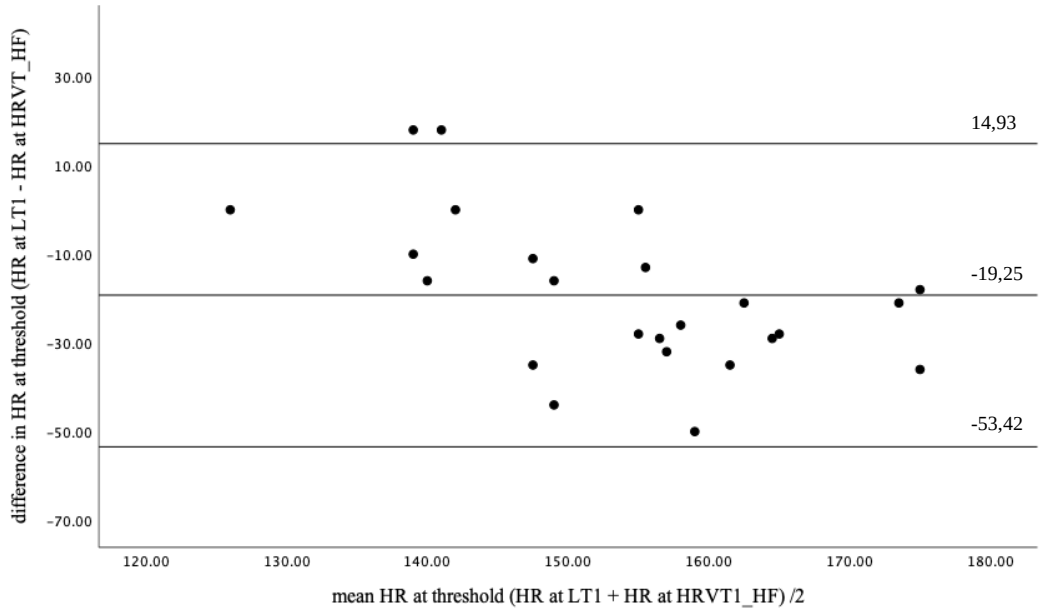
* $p < 0,05$, ** $p < 0,001$

When differences in HR at LT1 and HRVT1s were compared with Bland-Altman -plots, they showed smallest mean difference (-1,17 bpm) when HR at LT1 and HR at $HRVT1_{RMSSD}$ were compared. The second lowest difference was between HR at LT1 and $HRVT1_{SD1}$ (7,63 bpm). HR at $HRVT1_{HF}$ and at $HRVT1_{DFA}$ differed more from HR at LT1 (mean differences -19,25 and -20,58 bpm, correspondingly). The limits of agreement (LOA) in all comparisons were quite wide ($HRVT1_{RMSSD}$: ±30,25 bpm, $HRVT1_{HF}$: ±34,18 bpm, $HRVT1_{SD1}$: ±26,42 bpm, $HRVT1_{DFA}$: ±40,08 bpm). Bland-Altman plots of LT1–HRVT comparisons are presented in figure 12.

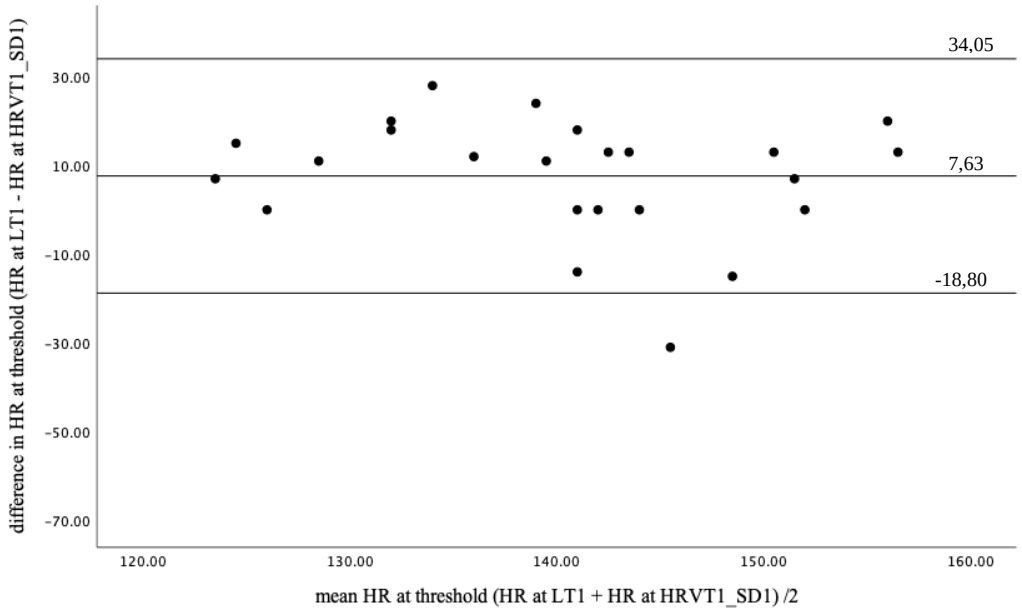
a)


29,08
-1,17
-31,42
30
10
-10
-30
-50
-70
120
130
140
150
160
170
difference in HR at threshold (HR at LT1 - HR at HRVT_RMSSD)
mean HR at threshold (HR at LT1 + HR at HRVT1_RMSSD)/2

b)


14,93
-19,25
-53,42
30.00
10.00
-10.00
-30.00
-50.00
-70.00
120.00
130.00
140.00
150.00
160.00
170.00
180.00
difference in HR at threshold (HR at LT1 - HR at HRVT_HF)
mean HR at threshold (HR at LT1 + HR at HRVT1_HF) /2

c)



d)

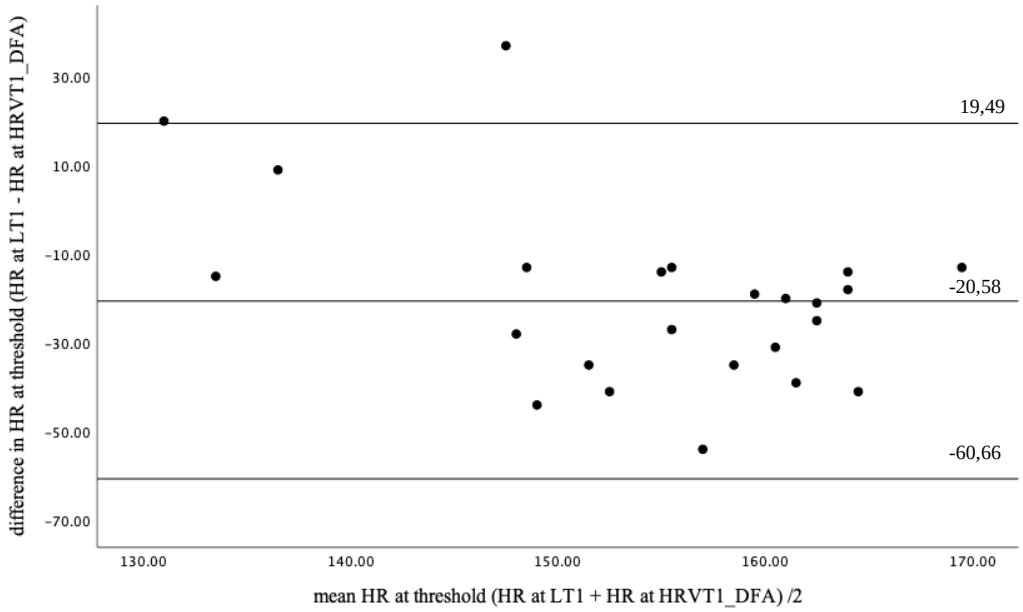
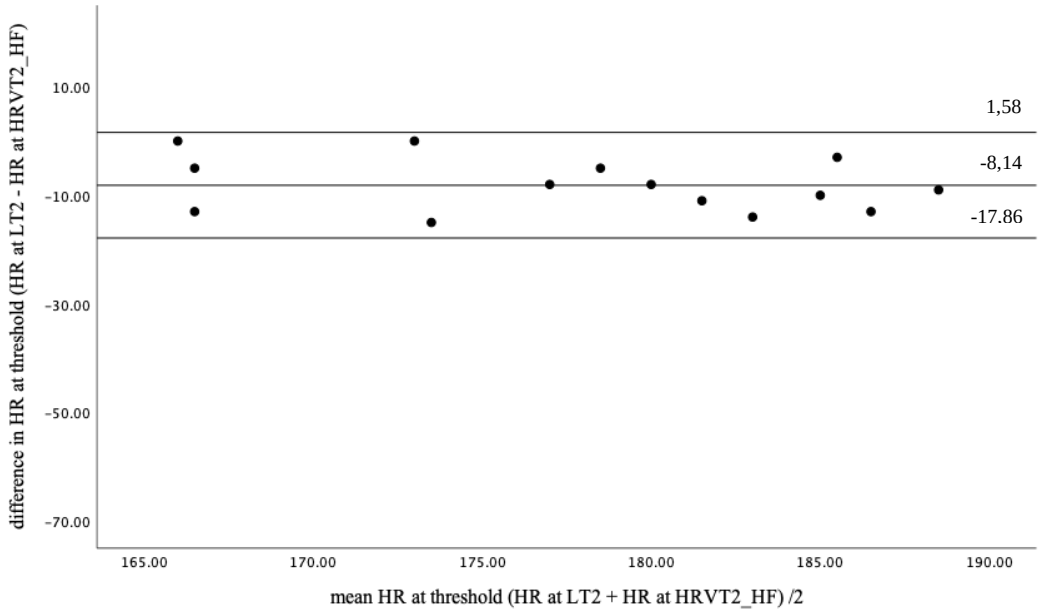


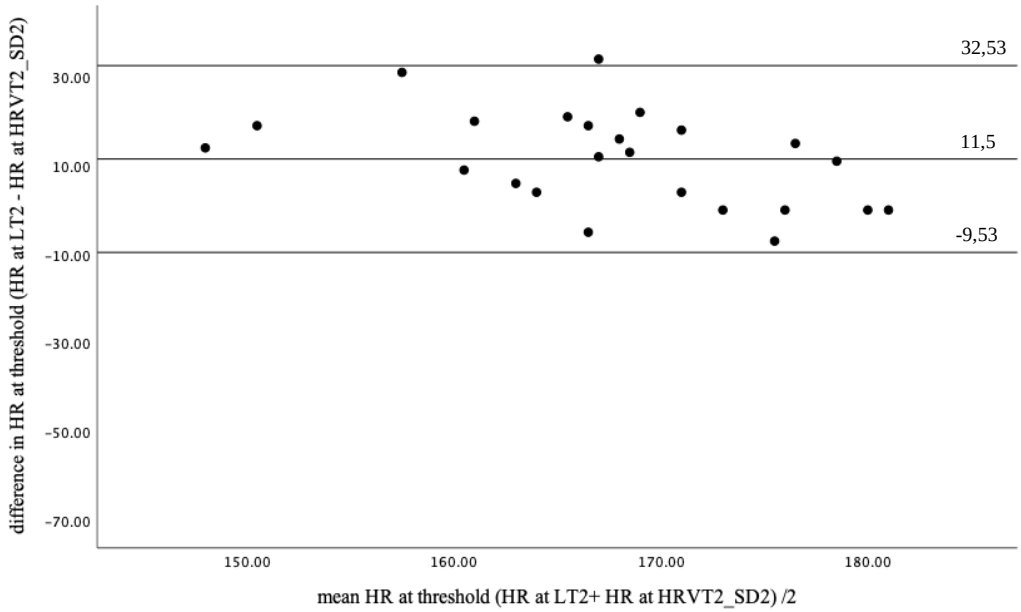
FIGURE 12. Bland Altman plots of the relationship in HR between LT1 and HRVT1 determined using RMSSD (a), HF (b), SD1 (c) and DFA-a1 (d) parameters.

Second threshold comparisons showed lowest mean difference in HR between LT2 & $HRVT2_{DFA}$ (mean difference -4,5 bpm). The differences in HR were smaller also between LT2 and $HRVT2_{HF}$ (-8,14 bpm) and between LT2 and $HRVT2_{SD2}$ (11,5 bpm) when compared to LT1 methods. The LOAs were narrowest when HR at LT2 and HR at $HRVT2_{HF}$ were compared (±9,72). Otherwise, they were quite wide ($HRVT2_{SD2}$: ±21,03 bpm, $HRVT2_{DFA}$: ±30,22 bpm). Bland-Altman plots of HR comparisons between LT2 and HRVT2s are shown in figure 13.

a)


difference in HR at threshold (HR at LT2 - HR at HRVT2_HF)
10.00
-10.00
-30.00
-50.00
-70.00
1,58
-8,14
-17.86
165.00
170.00
175.00
180.00
185.00
190.00
mean HR at threshold (HR at LT2 + HR at HRVT2_HF) /2

b)


difference in HR at threshold (HR at LT2 - HR at HRVT2_SD2)
30.00
10.00
-10.00
-30.00
-50.00
-70.00
32,53
11,5
-9,53
150.00
160.00
170.00
180.00
mean HR at threshold (HR at LT2+ HR at HRVT2_SD2) /2

c)

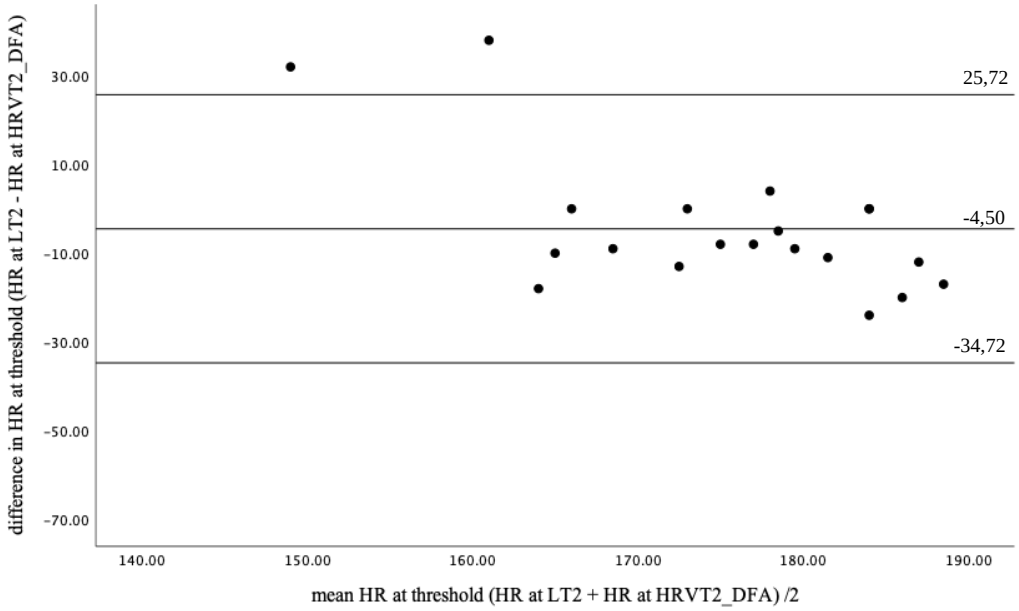
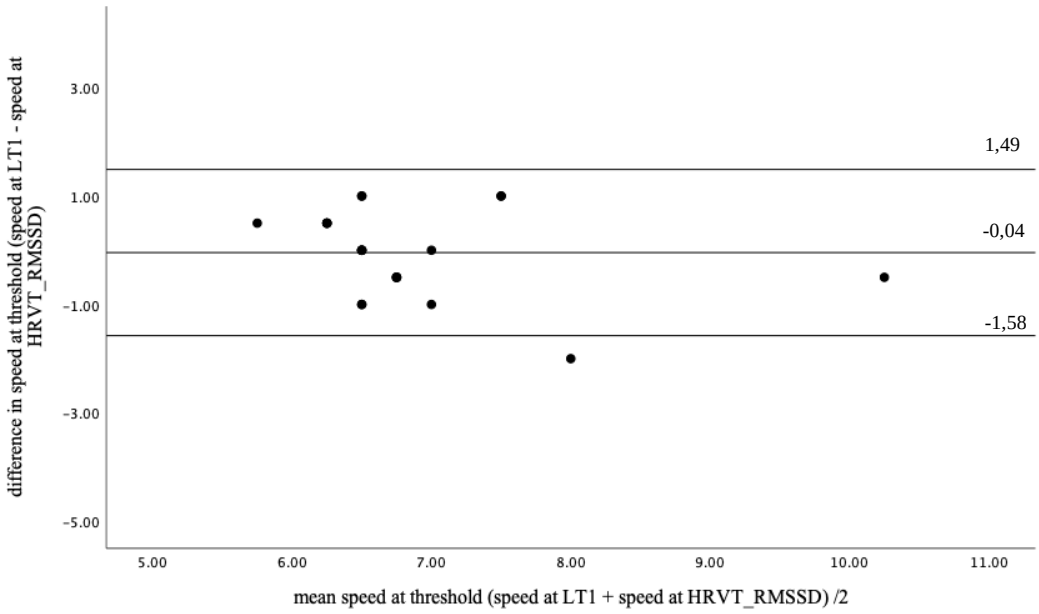


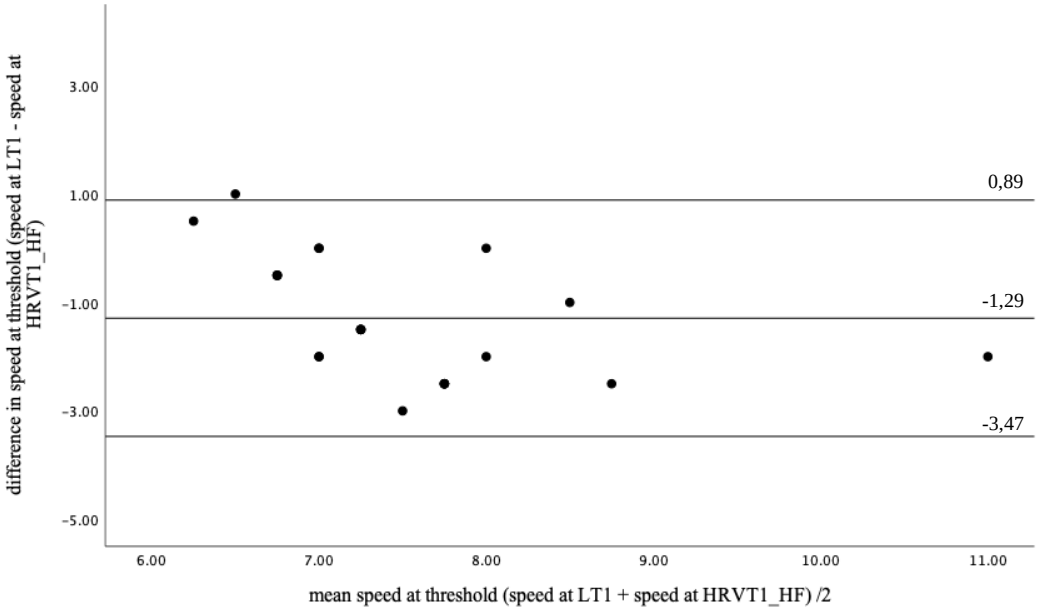
FIGURE 13. Bland Altman plots about the relationship in HR between LT2 and HRVT2 using HF (a), SD2 (b) and DFA-a1 (c) parameters.

When the differences in speed at LT1 and HRVT1s were depicted with Bland-Altman plots the smallest average differences were found between LT1 and $HRVT1_{RMSSD}$ (-0,04 km/h) and between LT1 and $HRVT1_{SD1}$ (0,40 km/h). The differences with two other HRVT methods ($HRVT1_{HF}$ and $HRVT1_{DFA}$) were more prominent (-1,29 km/h and -1,44 km/h). LOAs were also the narrowest when LT1 was compared with $HRVT1_{RMSSD}$ (±1,53 km/h) and with $HRVT1_{SD1}$ (±1,38 km/h). $HRVT1_{HF}$ and $HRVT1_{DFA}$ showed wider LOA (±2,18 km/h and ±2,72 km/h) in comparison with LT1 speed. Bland Altman plots of HRVT1 comparisons with LT1 speed are presented in figure 14.

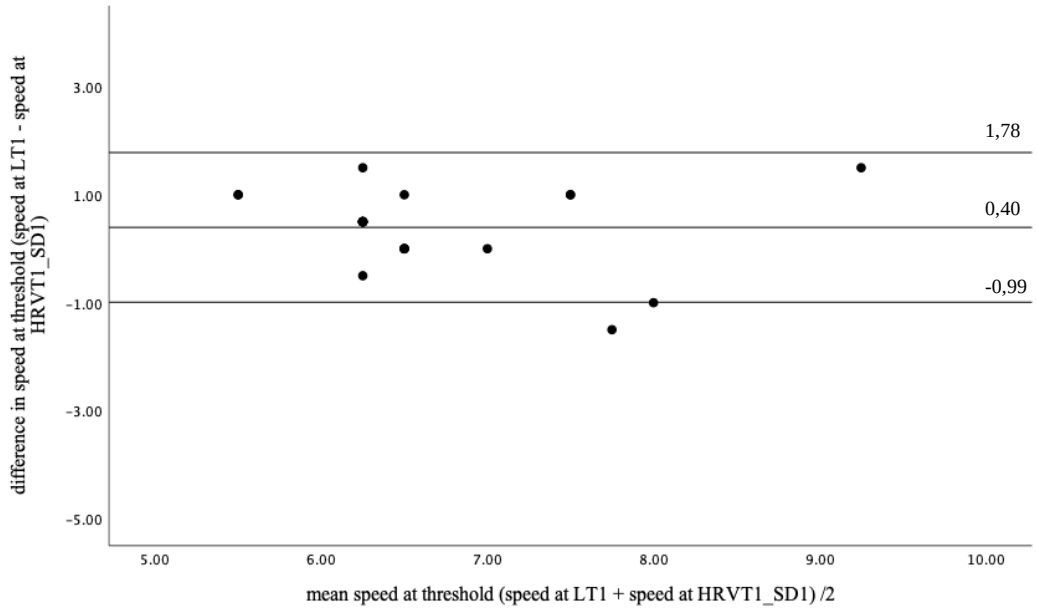
a)


difference in speed at threshold (speed at LT1 - speed at HRVT_RMSSD)
3.00
1.00
-1.00
-3.00
-5.00
1,49
-0,04
-1,58
5.00
6.00
7.00
8.00
9.00
10.00
11.00
mean speed at threshold (speed at LT1 + speed at HRVT_RMSSD) /2

b)


difference in speed at threshold (speed at LT1 - speed at HRVT1_HF)
3.00
1.00
-1.00
-3.00
-5.00
0,89
-1,29
-3,47
6.00
7.00
8.00
9.00
10.00
11.00
mean speed at threshold (speed at LT1 + speed at HRVT1_HF) /2

c)



d)

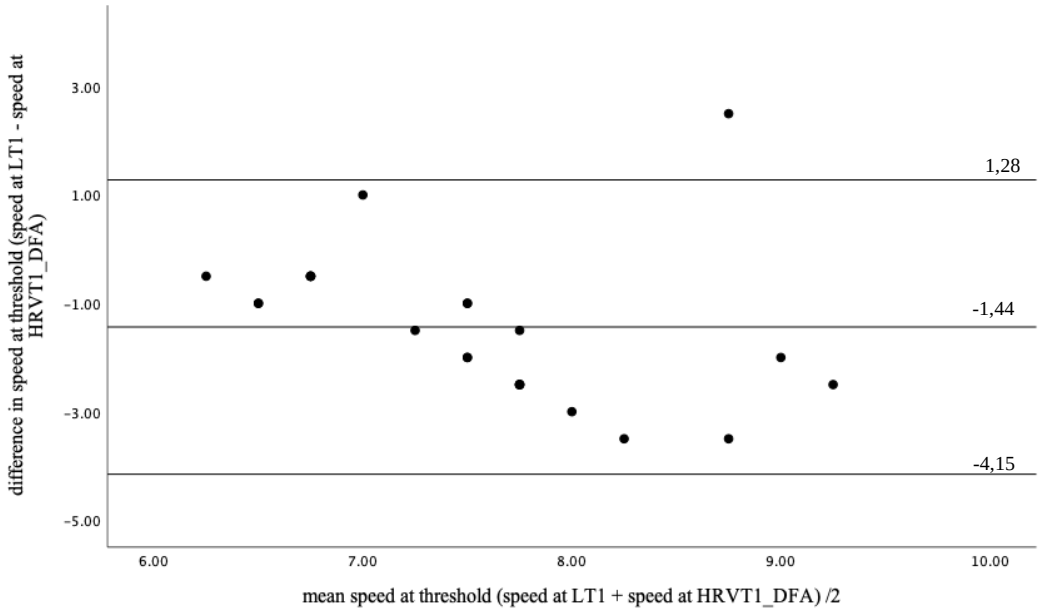
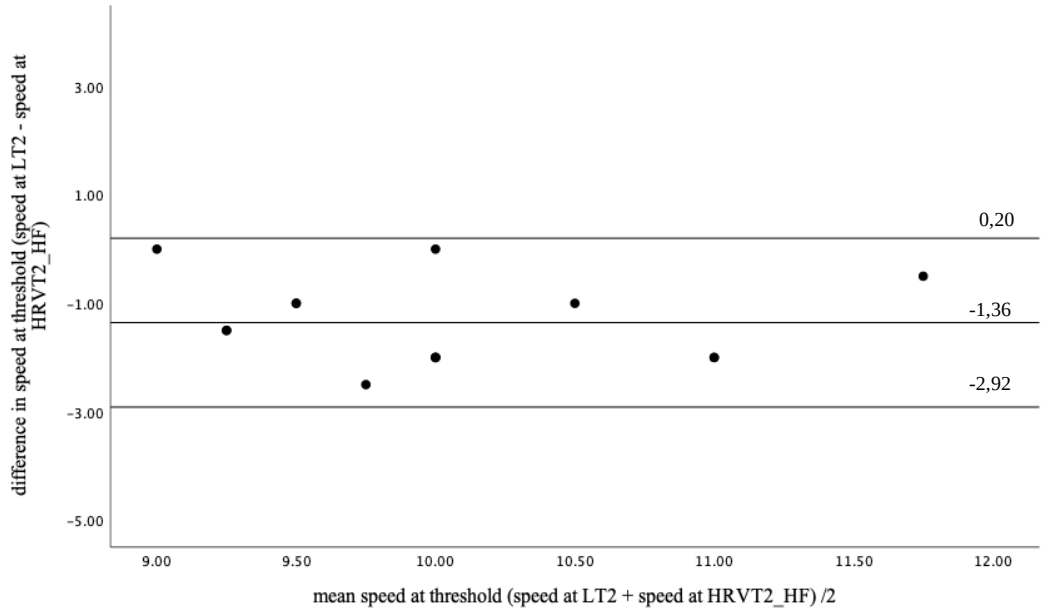


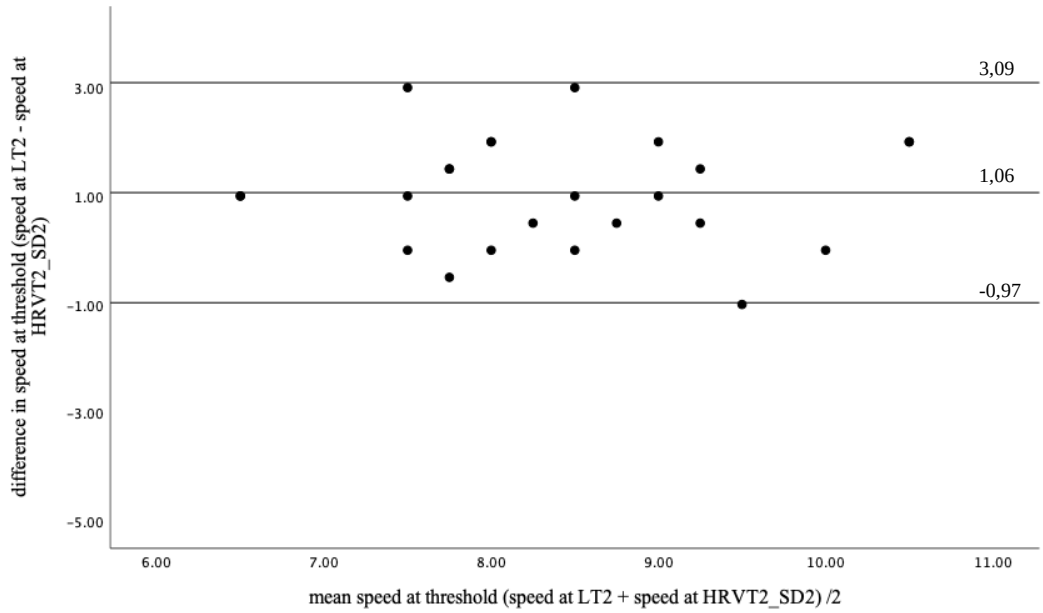
FIGURE 14. Bland Altman plots about the relationship in speed between LT1 and HRVT1 determined using RMSSD (a), HF (b), SD1 (c) and DFA-a1 (d) parameters.

When LT2 speed was put into Bland-Altman comparison with HRVT2 speeds the mean difference was smallest with $HRVT2_{DFA}$ method (-0,75 km/h). $HRVT2_{HF}$ and $HRVT2_{SD2}$ methods resulted in mean differences of -1,36 km/h and 1,06 km/h, correspondingly, when compared to LT2. LOAs were narrowest when LT2 was compared with $HRVT2_{HF}$ (±1,56 km/h), second lowest with $HRVT2_{SD2}$ (±2,03 km/h) and widest with $HRVT2_{DFA}$ (±3,12 km/h). Bland Altman plots of second threshold speed comparisons are presented in figure 15.

a)


difference in speed at threshold (speed at LT2 - speed at HRVT2_HF)
3.00
1.00
-1.00
-3.00
-5.00
0,20
-1,36
-2,92
9.00
9.50
10.00
10.50
11.00
11.50
12.00
mean speed at threshold (speed at LT2 + speed at HRVT2_HF) /2

b)


difference in speed at threshold (speed at LT2 - speed at HRVT2_SD2)
3.00
1.00
-1.00
-3.00
-5.00
3,09
1,06
-0,97
6.00
7.00
8.00
9.00
10.00
11.00
mean speed at threshold (speed at LT2 + speed at HRVT2_SD2) /2

c)

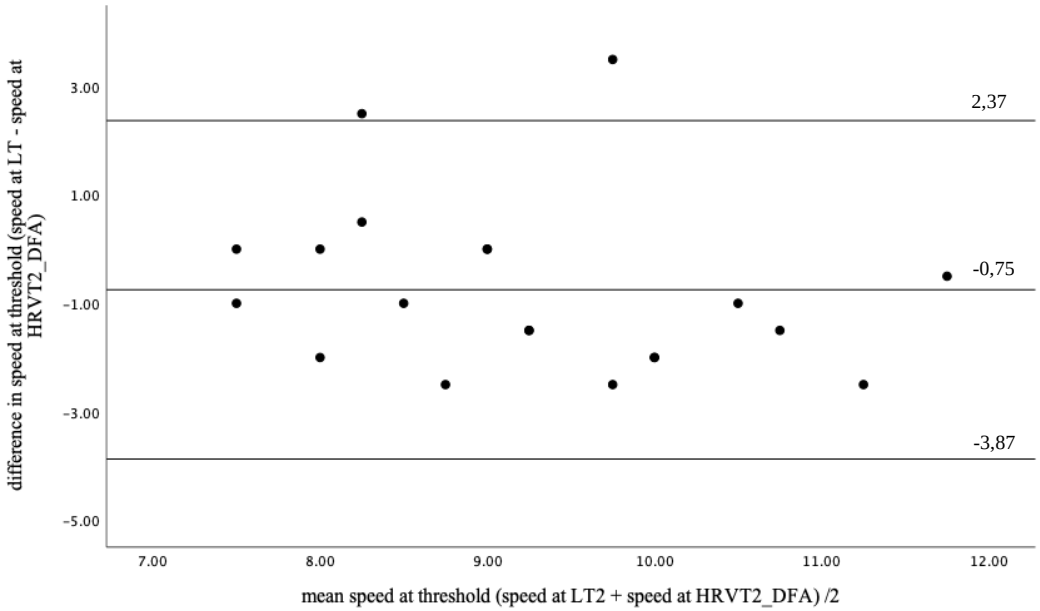


FIGURE 15. Bland Altman plots about the relationship in speed between LT2 and HRVT2 determined using HF (a), SD2 (b) and DFA-a1 (c) parameters.

The results of intraclass correlation coefficient (ICC) calculations are presented in tables 7 and 8. ICC revealed moderate reliability in threshold HR values only between $HRVT1_{RMSSD}$ and LT1 (0,519, CI: -0,161–0,798) and $HRVT2_{HF}$ and LT2 (0,708, CI: -0,227–0,927). Regarding threshold speed, moderate reliability was found between $HRVT2_{SD2}$ and LT2 (0,619, CI: -0,120–0,858) and between $HRVT2_{DFA}$ and LT2 (0,528, CI: -0,084–0,805). Highest ICC values and good reliability in threshold speed was found between $HRVT1_{RMSSD}$ (0,813, CI: 0,556–0,921) and LT1 and $HRVT1_{SD1}$ and LT1 (0,781, CI: 0,443–0,909).

TABLE 7. ICC estimates and 95 % confidence intervals of heart rate at LT and HRVTs.

HR	ICC	95% confidence intervals	
		Lower bound	Upper bound
$HRVT1_{RMSSD}$ & LT1	0,519	-0,161	0,798
$HRVT1_{HF}$ & LT1	0,322	-0,232	0,673
$HRVT1_{SD1}$ & LT1	0,436	-0,156	0,741
$HRVT1_{DFA}$ & LT1	-0,023	-0,431	0,384
$HRVT2_{HF}$ & LT2	0,708	-0,227	0,927
$HRVT2_{SD2}$ & LT2	0,411	-0,218	0,738
$HRVT2_{DFA}$ & LT2	0,428	-0,382	0,769

HR = heart rate, $HRVT1_{RMSSD}$ = first heart rate variability threshold determined from RMSSD, LT1 = first lactate threshold, $HRVT1_{HF}$ = first heart rate variability threshold determined from HF*fHF, $HRVT1_{SD1}$ = first heart rate variability threshold determined from SD1, $HRVT1_{DFA}$ = first heart rate variability threshold determined from DFA-a1, $HRVT2_{HF}$ = second heart rate variability threshold determined from HF*fHF, LT2 = second lactate threshold, $HRVT2_{SD2}$ = second heart rate variability threshold determined from SD2, $HRVT2_{DFA}$ = second heart rate variability threshold determined from DFA-a1, ICC = intraclass correlation coefficient

TABLE 8. ICC estimates and 95 % confidence intervals of speed at LT and HRVTs.

speed	ICC	95% confidence intervals	
		Lower bound	Upper bound
$HRVT1_{RMSSD}$ & LT1	0,813	0,556	0,921
$HRVT1_{HF}$ & LT1	0,469	-0,222	0,781
$HRVT1_{SD1}$ & LT1	0,781	0,443	0,909
$HRVT1_{DFA}$ & LT1	0,169	-0,272	0,540
$HRVT2_{HF}$ & LT2	0,429	-0,233	0,805
$HRVT2_{SD2}$ & LT2	0,619	-0,120	0,858
$HRVT2_{DFA}$ & LT2	0,528	-0,084	0,805

$HRVT1_{RMSSD}$ = first heart rate variability threshold determined from RMSSD, LT1 = first lactate threshold, $HRVT1_{HF}$ = first heart rate variability threshold determined from HF*fHF, $HRVT1_{SD1}$ = first heart rate variability threshold determined from SD1, $HRVT1_{DFA}$ = first heart rate variability threshold determined from DFA-a1, $HRVT2_{HF}$ = second heart rate variability threshold determined from HF*fHF, LT2 = second lactate threshold, $HRVT2_{SD2}$ = second heart rate variability threshold determined from SD2, $HRVT2_{DFA}$ = second heart rate variability threshold determined from DFA-a1, ICC = intraclass correlation coefficient

10 DISCUSSION

In the present study the applicability of heart rate variability thresholds was tested and compared with more traditional threshold method which is based on changes in blood lactate concentration. All the study participants were female novice runners. In almost all cross-sectional, intervention, and longitudinal studies dealing with HRV thresholds, and published so far, the participants have been males and mostly trained (Gronwald & Hoos 2019). Hence, the results of this study offer novel information and widen the knowledge about the applicability of HRVTs for fitness testing of different populations.

One of the main findings of this study was that both HRVTs were mostly detectable in the participants of this study. HRVT1 was detected by each method for every participant except by RMSSD-method for one participant of whom HRV did not show RMSSD values below threshold value (3 ms^2). $HRVT1_{SD1}$ has been detectable to a similar degree in previous study conducted for untrained individuals, too. In that same study $HRVT1_{RMSSD}$ was successfully determined for 60/68 of the participants which is slightly lower relative number than in the present study (88,2% vs. 95,8%). (Novelli et al. 2018) HRVT2 was not as comprehensively detectable among the participants. $HRVT2_{HF}$ could not be determined for 10 participants (42 %) and $HRVT2_{DFA}$ for 4 participants (17 %). Identification rate of $HRVT2_{HF}$ was lower than in previous studies where it has been successfully determined for all participants (Buccheit et al. 2007; Cottin et al. 2005; Cottin et al. 2006).

Correlations in HR and speed were analyzed using Spearman's rho values and linear regression graphs. Spearman's rho revealed the most significant correlations ($p < 0,001$) when speed in LT1 and $HRVT1_{SD1}$ and speed in LT2 and $HRVT2_{SD2}$ were compared (table 5). Less significant correlations ($p < 0,05$) were found for HR between LT1 and $HRVT1_{HF}$ as well as when LT2 was compared with $HRVT2_{HF}$ and with $HRVT2_{DFA}$. However, correlations are only showing the trends between changes of two variables. Additionally, even the highest correlation coefficients of the two pairs with lowest p-value stayed below 0,7, which is showing only moderately strong correlation (Chan 2003). For example, between LT1 and $HRVT1_{HF}$ average absolute HR values differed markedly as mean HR at $HRVT1_{HF}$ was 20 bpm higher than at LT1. Thus, it can be concluded that $HRVT1_{HF}$ is systemically overestimating HR. Linear regressions (appendix 2 & 3) showed the most significant, although moderate, relationship in HR between LT2 and $HRVT2_{HF}$ ($r = 0,7$) . Yet, this finding should be interpreted with caution

as $HRVT2_{HF}$ was determined for only half of the participants. Hence, the amount of data for this comparison was significantly smaller. Speed value comparisons revealed fair (r = 0,5) correlation between LT1 and $HRVT1_{SD1}$ and between LT1 and $HRVT1_{RMSSD}$. Regarding all other pairs, the coefficients of determination were low, and relationships were not found.

When the differences between HRVTs and LTs were compared, the smallest difference, and thus, the highest similarity, was found between LT1 and $HRVT1_{RMSSD}$. However, the standard errors were large for all threshold comparisons, including those two. SEs were the highest when $HRVT1_{DFA}$ was compared to LT1. Bland-Altman -plots supported these findings as they showed lowest mean difference between LT1 and $HRVT1_{RMSSD}$. LT1 and $HRVT1_{SD1}$ comparison showed also quite low difference. In all HR comparisons, the limits of agreement (LOA) were quite wide. LOAs in speed comparisons were narrower. When second threshold methods were compared, the lowest mean difference was found between LT2 and $HRVT2_{DFA}$. Differences were entirely lower when LT2 and HRVT2 methods were compared than LT1 comparisons. Narrowest LOAs were found when LT2 and $HRVT2_{HF}$ were compared. Otherwise, they were quite wide.

When intraclass correlation coefficients (ICC) and confidence intervals (CI) were analyzed, the only methods which showed moderate reliability as determinator of HR at LT were $HRVT1_{RMSSD}$ and $HRVT2_{HF}$. The reliability of HRVT-methods for defining threshold speed was better. Good reliability in comparison with LT1 speed was found between two methods: $HRVT1_{RMSSD}$ and $HRVT1_{SD1}$. Moderate reliability was also found when LT2 speed was compared with speed at $HRVT2_{SD2}$ and at $HRVT2_{DFA}$. It has been previously suggested that, in some cases, low ICC value can result from small subject number or lack of variability among measurement values instead of purely low agreement between measurements. Thus, it has been estimated that for reliable ICC comparisons at least 30 heterogenous samples would be needed. (Koo & Li 2016) In the present study 14–24 results depending on the variable were included to the ICC analysis which is lower amount than suggested. Yet, confidence intervals were also wide which is depicting the high interindividual variation in HRV results. One noteworthy thing is that the upper limit of confidence intervals in HR at $HRVT2_{HF}$ as well as in speed at $HRVT1_{RMSSD}$ and at $HRVT1_{SD1}$ showed excellent correlation ($CI_{upper\ limit} > 0,9$) compared to LT method, at least in some individuals. However, the low number of participants of whom $HRVT2_{HF}$ could be determined have possibly influenced the high ICC values that should be taken into account.

10.1 The applicability of heart rate variability thresholds

To conclude the results of different analysis, $HRVT1_{RMSSD}$ seems the most suitable method for determining first exercise threshold (ET) based on HRV measurements. Also, $HRVT1_{SD1}$ method showed moderate accuracy at least when correlations, Bland-Altman plots and ICCs are taken into consideration. Of HRVT2 methods used in this study no one seems significantly superior to others. Similar method for $HRVT1_{RMSSD}$ determination as used in the current study has been shown to correlate significantly ($p < 0,01$) with LT1 and VT1 in healthy middle-aged people and in same aged people with type 2 diabetes. Correlations of similar degree was also found for LT1 and $HRVT1_{SD1}$. (Sales et al. 2011) In the present study, Spearman's rho revealed significant correlations ($p < 0,001$) in speed between LT1 and $HRVT1_{SD1}$, too, while no correlations ($p > 0,05$) between LT1 and $HRVT1_{RMSSD}$ were found. Instead, paired differences analysis showed highest similarity between $HRVT1_{RMSSD}$ and LT1 of all analyzed pairs. In the previous study of Sales et al. (2011) the testing protocol was conducted with cycle ergometer and thresholds were only determined as cycling intensity (watts) hence results are not directly comparable to the results of this study. The participants were also mainly men. (Sales et al. 2011) In another previous study, a running test was conducted for male runners and $HRVT1_{SD1}$ speed was found correlate with LT1 speed ($p < 0,01$) while in that study the threshold determination was conducted in a slightly different way than in ours (Nascimento et al. 2017). In this study's novice female runners, Bland Altman -plots revealed even better agreement when $HRVT1_{SD1}$ was compared with LT1. Previous evidence also suggests that $HRVT1_{RMSSD}$ and $HRVT1_{SD1}$ may agree moderately with VT1 when some skiing techniques are used, although not all (Mendia-Iztueta et al. 2016). In our study $HRVT1_{SD1}$ tended slightly underestimate LT1 which has also been evident in a couple of previous studies (Mendia-Iztueta et al. 2016; Sales et al. 2011). The results of this study also suggest that $HRVT2_{SD2}$ method tends to underestimate LT2. This finding is not supported by the previous evidence where $HRVT2_{SD2}$ has been seen to match LT2 with low LOAs. However, similar correlations were found ($p < 0,001$) in previous studies. (Nascimento et al. 2017)

In the present study, HRVT determination method for first and second ET which was based on frequency analysis did not show strong applicability as it resulted in quite large variation. $HRVT2_{HF}$ was also impossible to determine in half of the participants. For some individuals it still might serve as accurate method for LT estimation as moderate relationship in HR ($r = 0,69$) was found between LT2 and $HRVT2_{HF}$ methods. Also, slight correlation ($p < 0,05$) in HR was

found when LT1 and $HRVT1_{HF}$ as well as LT2 and $HRVT2_{HF}$ were compared. In terms of speed, slight correlation ($p < 0,05$) was found for second threshold. $HRVT2_{HF}$ has been found to correlate with LT2 similarly in swimmers ($p < 0,05$) with no significant differences ($p > 0,05$) (Di Michele et al. 2012). Linear relationship of similar degree for second threshold has also been found in a study where trained young males were tested (Buccheit et al. 2007). However, in previous studies where HF-derived thresholds have been compared with VTs stronger correlations ($p < 0,001$) and smaller differences have usually been detected (Cottin et al. 2005; Cottin et al. 2006). Yet, for example in a study by Mendia-Iztueta et al. (2016) the effect of locomotion interfered the accurate estimation of VT2 by $HRVT2_{HF}$. The associations could have been better in their and in our study if the effect of locomotion on frequency components would have been considered, as suggested in some previous studies (Mourot et al. 2013). However, to this study, the method which only acknowledged the peak frequency was chosen. One possible reason why the correlations were weaker in the current study might be that novice runners' running technique pose more irregular rhythms that may affect more on frequency domain measures when compared to athletes that mostly have been tested in previous studies. Thus, it has been estimated that, for example, in cycling the pedaling motion accounts for 60 % of HRV independently of workload or HR (Blain et al. 2009). As already mentioned, in the present study all the participants were females which differed from previous studies where the methods have been tested mainly in men (Cottin et al. 2005; Cottin et al. 2006). Women have naturally greater high frequency variation (Ryan et al. 1994) which may be one reason why in this study HRVTs determined from HF-power showed higher average values than LTs.

In previous studies, strong relationships between threshold HRs have been found when $HRVT1_{DFA}$ and VT1 have been compared with Pearson's correlation coefficients, ICCs, and Bland Altman analysis (Rogers et al. 2021a). Also, HR and cycling power at LT1 and $HRVT1_{DFA}$ have been detected to correlate in different subject groups (although mainly men) when comparisons have been made using Pearson's r (Mateo-March et al. 2022; Rogers et al. 2021c; Rogers et al. 2022). These findings were not verified with the results of this study as $HRVT1_{DFA}$ showed quite significant inaccuracies when compared to LT1. Both threshold HR and speed differed in terms of correlations and paired differences analysis. LOAs were also the widest when compared to other HRVT1 methods. Though, Bland Altman plots have revealed quite wide LOAs between LTs and HRVTs also in previous studies (Mateo-March et al. 2022; Rogers et al. 2022). It should be noted that there are a couple of clear outliers in the Bland Altman plots of DFA-thresholds in the present study. Their HR at $HRVT1_{DFA}$ and $HRVT2_{DFA}$

was significantly lower than when determined by LT methods. Those results may partly widen the LOAs. Female sex may explain some of the differences detected in this study results compared to previous research results as it has been suggested that women have more complex HR dynamics than men (Ryan et al. 1994). Hence, it is likely that this is not the only explanation for the observed differences.

Similar to this study, threshold pace has showed better agreement than HR in previous studies where LTs and DFA-thresholds have been compared (Mateo-March et al. 2022; Rogers et al. 2022). For example, in a study by Mateo-March et al. (2022) $HRVT1_{DFA}$ power output showed correlation with LT1 but not HR. According to the current results DFA-a1 based methods do not seem to offer more accurate information about changes in physiological requirements during high intensity exercise, when compared to traditional time domain methods. Even though, e.g., in a study conducted for young elite male athletes that has been suggested (Blasco-Lafarga et al. 2017). Both DFA-a1 derived HRVTs seemed to overestimate LTs. The reasons for that are unclear, but the issues related to female sex and cardiolocomotor coupling may influence these detected differences compared to previous studies. In addition, the inaccuracies might be partly caused by Kubios artefact correction rates. With high exercise intensities as DFA-a1 values are declining, the correction bias substantially increases, which may result in false high DFA-a1 values. (Rogers et al. 2021d) However, decrease in DFA-a1 values with increasing exercise intensity was visible in all participants of this study, to some extent, which possibly reflected the changes in ANS-activity, in other words, increased SNS and reduced PNS activity (Hautala et al. 2003).

Mateo-March et al. (2022) found out that $HRVT1_{DFA}$ was better comparable to LT1 than $HRVT2_{DFA}$ to LT2 as both HR and power output at $HRVT2_{DFA}$ tended to underestimate LT2 in cycling GXT. Similar effect was not evident in our study as $HRVT2_{DFA}$ rather slightly overestimated HR and running speed at LT2. In another study by Rogers et al. (2021b) they suggested that HR at $HRVT2_{DFA}$ was closely correlated ($p < 0{,}001$) with HR at VT2 in recreational male runners with different fitness levels. When compared to results of our study it seems that the same method for second threshold determination is weaker estimator for LT2 in female novice runners. Question remains if the difference is caused by differences between VT and LT levels, between males and females, between test modes or some other issue or, maybe most possibly, the combination of those. Data regarding the effect of exercise duration on detrended fluctuation and the reliability of DFA-a1 method during high intensity exercise is

still weak and those issues may also possibly explain the differences between studies (Gronwald et al. 2019).

The results revealed quite remarkable differences in lactate concentrations when LT1 and HRVT1s were compared (table 2 & 3). For example, the mean lactate concentration at $HRVT1_{HF}$ and at $HRVT1_{DFA}$ were over 2-fold higher when compared to LT1 (2,8/2,9 mmol/l vs. 1,4 mmol/l). The same effect is evident when the mean lactate concentration at $HRVT2_{HF}$ was compared with LT2 (5,3 mmol/l vs. 3,3 mmol/l). In practice, the error in training zone boundary determination of that degree may result in negative training responses and interfere the reaching of wanted training stimuli (Gronwald et al. 2020). The differences between lactate concentrations at LT and HRVT illustrate that the physiological and biochemical changes at threshold intensities are probably significantly different (Gladden 2004; Robergs et al. 2004). Also, mean HR at $HRVT1_{HF}$ (164 bpm) and at $HRVT1_{DFA}$ (165 bpm) were both almost 20 bpm higher than with reference method (HR at LT1: 144 bpm) which is quite a huge difference.

A couple of individual deflections can be made from the results. Firstly, in some individuals the HRVTs assessed with different methods differed quite remarkably from each other, even though they should be depicting the same physiological changes. For example, in one individual speed in $HRVT1_{DFA}$ was 3 km/h lower than at $HRVT1_{RMSSD}$. In previous studies, it has been suggested that $HRVT1_{DFA}$ -method would be more accurate than methods based on traditional HRV-analysis as it is less prone to bias in the bout of exercise (Gronwald et al. 2019; Gronwald et al. 2020). Our findings do not support that as speed at $HRVT1_{DFA}$ was 1,5 km/h higher compared to LT1 and showed the most remarkable difference of all HRVT1 methods tested. Also, from Bland-Altman plots an interesting observation was made. Two methods, $HRVT1_{HF}$ and $HRVT1_{DFA}$ seemed to overestimate HR more clearly in those individuals with higher HR at LT1 (figure 12). The same trend can be slightly seen with HR at $HRVT2_{DFA}$ (figure 13) and also with speed at $HRVT1_{DFA}$ (figure 14). Reasons behind those differences are unclear. Instead, in a study by Cottin et al. (2006) they found that $HRVT1_{HF}$ method underestimated running speed more significantly in individuals with higher speed at VT1. In the current study, maybe the higher LT1 is reflecting higher individual fitness level. Individuals with higher fitness usually have higher HRV at certain absolute exercise intensity (Gronwald et al. 2020). Also, the overall HRV of adult females is decreasing slower when compared to males (Acharya et al. 2006). This may affect the comparisons with the previous research results because the threshold values used in this study are developed according to studies conducted mainly in men. Another

possible reason may be associated with cardiolocomotor coupling which might be more evident in individuals with higher fitness level and higher threshold levels. Some ambiguities in chosen threshold assessment criteria were also noted. For instance, in one subject RMSSD values clearly stabilized at about 3 ms^2 but did not clearly lower below that. The first threshold was still determined to that point where the stabilization started. In another subject, RMSSD did not lower below 3 ms^2 at all. For that individual, $HRVT1_{RMSSD}$ was not determined. The HRV behavior of that same individual was interesting also because at the same time RMSSD did not show any clear reduction the HRV measured by non-linear parameters reduced quite quickly. This is also one example of between methods differences detected in the present study.

10.2 Strengths of the study

The strengths of this study include that HRV was measured using Firstbeat Bodyguard 2 which has accuracy of same level than clinical use ECG (Parak & Korhonen s.a.). Thus, the precision of HRV measurements can be appraised to be superior when compared to basic HR monitor which has been used in many previous similar studies (Rogers et al. 2021d). The electrode placement was conducted as carefully as possible to minimize the possibility of electrode movements that would have induced the possibility for artefacts (Giles & Draper 2018). Additionally, the treadmill GXT protocol was carefully conducted according to Finnish guidelines and previous reliability and validity studies (Nummela & Peltonen 2018; Weltman et al. 1990).

Other significant strength of our study was that the number of participants was approximately two times higher than in many previous studies. In addition, all the participants were females. In a couple of previous studies examining the applicability of HRVTs it has been stated that there are very few studies where DFA-a1 is investigated in women and further studies to this topic in female population are needed (Rogers et al. 2021c; Rogers et al. 2022). In most of the studies investigating the validity of HRVT methods, subjects have been men with similar age, weight, and fitness levels (Cottin et al. 2006; Rogers et al. 2021a). Also, elite endurance athletes have been studied and they have been mainly men, as well. In previous studies, the test has often been cycling test instead of running, thus, the results of our study offer novel information about the applicability of HRVT methods in treadmill test. Whereas that also means that results are not totally comparable. (Mateo-March et al. 2022; Rogers et al. 2022)

Another strength of this study is that for HRVT determination multiple methods were used. The results were also analyzed using various statistical analysis. That is important as different methods pose different bias sources. For example, Bland Altman -plots method was chosen for the validity analysis as it is suggested to give more accurate information about the equality relationship of two methods than, for instance, linear regressions (Bland & Altman 1999). However, that statement has later been questioned. It has been proposed that Bland Altman plots and LOA may give false information about the validity of a new method if the compared traditional method has imprecise nature and/or the methods have different level of bias. (Hopkins et al. 2008) Hence, when the final conclusions have been conducted as a combination of various analysis methods, the effect of possible inaccuracies in one method is suggested to be minor.

10.3 Limitations

There were also certain limitations in the present study. Firstly, the treadmill was stopped for about 10 s every time when the fingertip blood sample was taken. That stopping was visible in the HRV recording curves as reduction or flattening of the curve. Thus, HRV either decreased or stopped linear increasing for a little while when running was stopped, and blood sample was taken. These short pauses in running between stages may affect the results when compared to previous studies, which have been conducted in a continuous manner. (Buccheit et al. 2007; Cottin et al. 2006; Nascimento et al. 2017; Rogers et al. 2021a & 2021b). HRV is prone to bias especially during exercise hence the sampling pauses may have affected autonomic modulation and, hence, be seen in HRV (Sandercock & Brodie 2006). However, the time phases from which HRV was measured did not overlap those moments when blood sampling was conducted. Instead, the HRV was calculated at the end of each test stage just before treadmill stopping, thus, the effect of that was supposed to be minor. Additionally, previous studies have suggested that HRV recovery (i.e., increase following acute exercise) is faster the lower the preceding exercise intensity has been, and as regards to high intensity exercise, the acute recovery is very faint (Michael et al. 2016; Stanley et al. 2013). Thus, the possible minor effect of pauses in running may have an influence only in the stages with the lowest intensity.

Another possible limitation in the study protocol was that the threshold assessment criteria was chosen according to previous studies but the reproducibility of the methods in novice female runners was not tested. Yet, reproducibility of some of the methods used has been tested previously and, for example, the 3 ms^2 threshold method for $HRVT1_{RMSSD}$ and $HRVT1_{SD1}$ has been shown moderate to high reproducibility in untrained, young subjects (Novelli et al. 2018). The reproducibility of lactate thresholds should have also been tested in that particular study population as there might be inaccuracy also in "gold standard" -lactate threshold determinations. Previous studies have shown that the reproducibility of LTs especially in untrained individuals may be poor. For example, in a study by Grant et al. (2002) they found weak correlation coefficients for LTs in untrained compared to trained men.

In our study, the accurate assessment of lactate thresholds turns out to be difficult in some individuals thus it may have led to imprecision in reference thresholds. One reason might be that the participants in the present study had lower fitness level when compared to previous studies. The lactate profile of untrained individuals is generally different compared to trained individuals because lactate appearance in blood at especially submaximal training intensities is decreased because of endurance training. (MacRae et al. 1992) Additionally, bias level between lactate and HRV measurements may have possibly been different as HRV is quite prone to bias especially during exercise (Sandercock & Brodie 2006). As already mentioned, repeatability of HRVT nor LT methods was not tested which may have affected, for example, to Bland Altman -results. Bland and Altman (1999) recommend testing the repeatability of compared methods because poor agreement between methods may in some cases also result from poor repeatability of one or both methods. Because of aforementioned issues conclusions from the Bland Altman plots per se cannot be made without cautiousness.

One consideration regarding the interpretation of the results is that there might be considerable interindividual variation in resting levels of HRV parameters. That may cause differences and inaccuracies to thresholds which are determined to a certain fixed value ($HRVT1_{RMSSD}$, $HRVT1_{SD1}$ and $HRVT2_{SD2}$). The variation is resulting from multiple different properties and conditions affecting HRV, e.g., age, fitness level and disease status (Shaffer et al. 2014). Hence, even if the absolute threshold value of the analyzed parameter would be the same, marked differences may emerge if relative differences to resting levels are compared between individuals. The question whether the differences in resting levels of those variables affected the determined HRVTs or not was beyond the framework of this study and was not tested.

However, that is one issue which would be interesting to be investigated and considered in future studies.

As already mentioned, HRV is prone to bias especially during quickly changing physiological conditions that exercise usually is (Sandercock & Brodie 2006). There was notable bias in HRV recordings of some individuals which may have affected HRV values taken into analysis. To minimize the amount of bias, the recording phases for the final analysis were carefully chosen and if marked false beats occurred during the last 60 s frame, from which the HRV parameters were calculated, the analysis window was either shortened or moved to an earlier time of that stage. There was also other possible bias in the trends of HRV parameters as e.g., DFA-a1 values did not show direct linear trend in all individuals. Instead, in some subjects the values were constantly changing back and forth during GXT. Also, the effect of stride frequency on HRV (especially in high exercise intensities) is known but it was not considered in threshold analysis (Di Michele et al. 2012; Cottin et al. 2004; Gronwald & Hoos 2019). According to previous studies, at least in cycling, the effect of cardio locomotor coupling is more markedly evident the higher is exercise intensity (Blain et al. 2009). That is also one reason why the capacity of HRV methods, based on spectral analysis, to capture real changes in HRV may be limited. It is also known that in high exercise intensities also other factors than autonomic control mechanisms have effect on HRV. (Sandercock & Brodie 2006) Yet the effects of those factors are hard to capture.

Considering that all the subjects were females, one limitation of the study was that menstrual cycle phase was not controlled or reported. Menstrual cycle phase in the test day could possibly have had some individual effects on measured physiological markers, especially on HRV. The influence is yet considered to be faint as evidence from previous studies suggest that menstrual cycle may only have minor, individual effects on performance variables or thresholds measured during GXT. However, the research data about that topic is limited. (Smekal et al. 2007; Taipale-Mikkonen et al. 2021) Additionally, nutrition nor hydration preceding the test was controlled and, for example, differences in carbohydrate intake may have affected slightly on measured lactate concentration (Aunola & Rusko 1984). Also, day to day differences occur in resting blood lactate, which may affect LT1 (Pfitzinger & Freedson 1998). It is unclear if those differences have some effect on HRV, too.

Some testing equipment and analysis program-related sources of bias also exist. In the current study, the artifact correction method level was kept at 3 %. It has been stated in previous studies that artifact rate of even 5% in high intensities may result in substantial errors despite correction method used. When exercise intensity is rising, so does the possibility for artifacts. That may blunt or drop HRV correlation properties. (Gronwald et al. 2020) However, even if the artefacts or their correction may cause bias into nonlinear HRV-parameters (e.g., DFA-a1) in high exercise intensities, it has been suggested that that would have no effect on HRVT determination (Rogers et al. 2021d).

10.4 Future consideration and applications

The results of this study suggest that first lactate threshold can be identified by measuring HRV of novice female runners. The most accurate HRVT method for defining LT1 in that population seemed to be the one based on traditional time domain analysis. With respect to second threshold determination, the HRVT methods tested here seemed inaccurate as none of the methods showed both strong correlations and high identifiability.

As stated already in multiple previous studies, in the future, it would be important to define and take into consideration the physiological factors that affect ANS regulation and HRV even more accurately. Those can be, for instance, cardiovascular drift, hydration status, emotional stress, temperature, air flow or anxiety level. General relationship between different HRV parameters and HR should also be investigated, as well as effects of sex, age, fitness level and cardiac risk factors. It is also important to bear in mind that laboratory settings where HRVT, as well as LT methods, are mainly tested, are different to field conditions where actual training is conducted. (Gronwald et al. 2020) Therefore the applicability of HRVTs for training prescription in practice cannot be evaluated based on these results.

With respect to HRVT methods which are based on fixed threshold values (RMSSD, SD1, SD2 and DFA-a1) it is also of interest if there is an individual threshold stages or separate threshold values depending on the study participants' characteristics (e.g., age, sex, fitness level), instead of fixed values. For example, according to the results of this study, $HRVT1_{DFA}$ seemed to quite significantly overestimate LT1, thus, it would be supposed that threshold value of above 0,75 would have resulted more accurate threshold estimation in this kind of female population.

The determination of physiological thresholds is always, to some degree, subjective and imprecise which should be bear in mind when interpreting the results. Some inaccuracy sources exist in the lactate threshold methods which were used here as gold standards. In most previous studies where the reliability and validity of physiological threshold methods have been tested lactate threshold methods have been compared to each other or to ventilatory threshold methods and vice versa. Those physiological parameters thus attribute certain bias sources per se, which may reduce the reliability. Hence, the same problem exists with HRVT methods, and the most relevant question is, what are the actual physiological functions wanted to be captured. The physiology behind endurance exercise is complex and there are many factors influencing it thus the assessment of e.g., physiological condition or exercise response by using only single measure is practically impossible (Gronwald et al. 2020). Maybe the most accurate estimates of individual ETs are captured by the combination of various physiological measures. The adjustment of threshold states is also important if desired training responses are not reached.

REFERENCES

Acharya R., Joseph K. P., Kannathal N., Lim C. M. & Suri J. S. (2006). Heart rate variability: a review. Medical & biological engineering & computing 44 (12), 1031–1051. doi:10.1007/s11517-006-0119-0.

Ahmaidi, S., Hardy, J. M., Varray, A., Collomp, K., Mercier, J. & Préfaut, C. (1993). Respiratory gas exchange indices used to detect the blood lactate accumulation threshold during an incremental exercise test in young athletes. European journal of applied physiology and occupational physiology 66 (1), 31–36. doi:10.1007/BF00863396.

Aunola, S. & Rusko, H. (1984). Reproducibility of aerobic and anaerobic thresholds in 20–50-year-old men. European journal of applied physiology and occupational physiology, 53 (3), 260–266. doi:10.1007/BF00776600.

Baron, B., Noakes, T. D., Dekerle, J., Moullan, F., Robin, S., Matran, R. & Pelayo, P. (2008). Why does exercise terminate at the maximal lactate steady state intensity? British journal of sports medicine, 42 (10), 828–833. doi:10.1136/bjsm.2007.040444.

Bergstrom, H. C., Housh, T. J., Zuniga, J. M., Traylor, D. A., Camic, C. L., Lewis, R. W., Jr, Schmidt, R. J. & Johnson, G. O. (2013). The relationships among critical power determined from a 3-min all-out test, respiratory compensation point, gas exchange threshold, and ventilatory threshold. Research quarterly for exercise and sport, 84 (2), 232–238. doi:10.1080/02701367.2013.784723.

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 (5), H1651–H1659. doi:10.1152/ajpheart.00881.2008.

Bland, J. M. & Altman, D. G. (1999). Measuring agreement in method comparison studies. Statistical methods in medical research 8 (2), 135–160. doi:10.1177/096228029900800204.

Blasco-Lafarga, C., Camarena, B. & Mateo-March, M. (2017). Cardiovascular and Autonomic Responses to a Maximal Exercise Test in Elite Youngsters. International Journal of Sports Medicine, 38 (9), 666–674. doi:10.1055/s-0043-110680.

Braun, C., Kowallik, P., Freking, A., Hadeler, D., Kniffki, K. D. & Meesmann, M. (1998). Demonstration of nonlinear components in heart rate variability of healthy persons. The

American journal of physiology 275 (5), H1577–H1584. doi:10.1152/ajpheart.1998.275.5.H1577.

Buccelletti, F., Bocci, M., Gilardi, E., Fiore, V., Calcinaro S., Fragnoli, C., Maviglia, R. & Franceschi, F. (2012). Linear and nonlinear heart rate variability indexes in clinical practice. Computational and mathematical methods in medicine, 2012, 219080. doi:10.1155/2012/219080.

Buchheit, M., Solano, R. & Millet, G. P. (2007). Heart-rate deflection point and the second heart-rate variability threshold during running exercise in trained boys. Pediatric exercise science 19 (2), 192–204. doi:10.1123/pes.19.2.192.

Cambri, L., Novelli, F., Sales, M., Sousa, L., Queiroz, M., Dias, A., Santos, K. & Arsa, G. (2016). Heart rate inflection point estimates the anaerobic threshold in overweight and obese young adults. Sport Sciences for Health 12, 397–405. doi:10.1007/s11332-016-0304-y.

Casties, J. F., Mottet, D. & Le Gallais, D. (2006). Non-linear analyses of heart rate variability during heavy exercise and recovery in cyclists. International journal of sports medicine 27 (10), 780–785. doi:10.1055/s-2005-872968.

Cerezuela-Espejo, V., Courel-Ibáñez, J., Morán-Navarro, R., Martínez-Cava, A. & Pallarés, J. G. (2018). The Relationship Between Lactate and Ventilatory Thresholds in Runners: Validity and Reliability of Exercise Test Performance Parameters. Frontiers in physiology 9, 1320. doi:10.3389/fphys.2018.01320.

Chan Y. H. (2003). Biostatistics 104: correlational analysis. Singapore medical journal 44 (12), 614–619.

Cottin, F., Médigue, C., Leprêtre, P. M., Papelier, Y., Koralsztein, J. P. & Billat, V. (2004). Heart rate variability during exercise performed below and above ventilatory threshold. Medicine and science in sports and exercise 36 (4), 594–600. doi:10.1249/01.mss.0000121982.14718.2a.

Cottin, F., Médigue, C., Lopes, P., Leprêtre, P. M., Heubert, R. & Billat, V. (2006). Ventilatory thresholds assessment from heart rate variability during an incremental exhaustive running test. International journal of sports medicine, 28 (4), 287–294. doi:10.1055/s-2006-924355.

Cox, J. H., Cortright, R. N., Dohm, G. L. & Houmard, J. A. (1999). Effect of aging on response to exercise training in humans: skeletal muscle GLUT-4 and insulin sensitivity. Journal of applied physiology 86 (6), 2019–2025. doi:10.1152/jappl.1999.86.6.2019.

Diaz-Canestro, C. & Montero, D. (2019). Sex Dimorphism of VO2max Trainability: A Systematic Review and Meta-analysis. Sports medicine (Auckland, N.Z.), 49 (12), 1949–1956. doi:10.1007/s40279-019-01180-z.

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 research, 26 (11), 3059–3066. doi:10.1519/JSC.0b013e318245bde1.

Durnin, J. V. & Womersley, J. (1974). Body fat assessed from total body density and its estimation from skinfold thickness: measurements on 481 men and women aged from 16 to 72 years. The British journal of nutrition, 32 (1), 77–97. doi:10.1079/bjn19740060.

Faude, O., Kindermann, W. & Meyer, T. (2009). Lactate threshold concepts: how valid are they? Sports medicine (Auckland, N.Z.) 39 (6), 469–490. doi:10.2165/00007256-200939060-00003.

Gaskill, S. E., Ruby, B. C., Walker, A. J., Sanchez, O. A., Serfass, R. C. & Leon, A. S. (2001). Validity and reliability of combining three methods to determine ventilatory threshold. Medicine and science in sports and exercise 33 (11), 1841–1848. doi:10.1097/00005768-200111000-00007.

Giles, D. A. & Draper, N. (2018). Heart Rate Variability During Exercise: A Comparison of Artefact Correction Methods. Journal of strength and conditioning research, 32 (3), 726–735. doi:10.1519/JSC.0000000000001800.

Gladden L. B. (2004). Lactate metabolism: a new paradigm for the third millennium. The Journal of physiology 558 (1), 5–30. doi:10.1113/jphysiol.2003.058701.

Goldberger, J.J., Le, F.K., Lahiri, M., Kannankeril, P.J., Ng, J. & Kadish, A.H. (2006). Assessment of parasympathetic reactivation after exercise. American Journal of Physiology, Heart and Circulatory Physiology 290 (6), H2446–H2452. doi: 10.1152/ajpheart.01118.2005.

Grant, S., McMillan, K., Newell, J., Wood, L., Keatley, S., Simpson, D., Leslie, K. & Fairlie-Clark, S. (2002). Reproducibility of the blood lactate threshold, 4 mmol/l marker, heart rate and ratings of perceived exertion during incremental treadmill exercise in humans. European journal of applied physiology, 87 (2), 159–166. doi:10.1007/s00421-002-0608-2.

Gronwald, T. & Hoos, O. (2019). Correlation properties of heart rate variability during endurance exercise: A systematic review. Annals of noninvasive electrocardiology: the

official journal of the International Society for Holter and Noninvasive Electrocardiology Inc 25 (1), e12697. doi:10.1111/anec.12697.

Gronwald, T., Rogers, B. & Hoos, O. (2020). Fractal Correlation Properties of Heart Rate Variability: A New Biomarker for Intensity Distribution in Endurance Exercise and Training Prescription? Frontiers in physiology 11, 550572. doi:10.3389/fphys.2020.550572.

Gronwald, T., Berk, S., Altini, M., Mourot, L., Hoos, O. & Rogers, B. (2021). Real-Time Estimation of Aerobic Threshold and Exercise Intensity Distribution Using Fractal Correlation Properties of Heart Rate Variability: A Single-Case Field Application in a Former Olympic Triathlete. Frontiers in sports and active living, 3, 668812. doi:10.3389/fspor.2021.668812.

Hardstone, R., Poil, S. S., Schiavone, G., Jansen, R., Nikulin, V. V., Mansvelder, H. D. & Linkenkaer-Hansen, K. (2012). Detrended fluctuation analysis: a scale-free view on neuronal oscillations. Frontiers in physiology, 3, 450. doi:10.3389/fphys.2012.00450.

Hautala, A. J., Kiviniemi, A. M., Mäkikallio, T. H., Kinnunen, H., Nissilä, S., Huikuri, H. V. & Tulppo, M. P. (2006). Individual differences in the responses to endurance and resistance training. European journal of applied physiology 96 (5), 535–542. doi:10.1007/s00421-005-0116-2.

Hautala, A. J., Mäkikallio, T. H., Seppänen, T., Huikuri, H. V. & Tulppo, M. P. (2003). Short-term correlation properties of R-R interval dynamics at different exercise intensity levels. Clinical physiology and functional imaging 23 (4), 215–223. doi:10.1046/j.1475-097x.2003.00499.x.

Hopkins, W. G., Marshall, S. W., Batterham, A. M. & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. Medicine and science in sports and exercise 41 (1), 3–13. doi:10.1249/MSS.0b013e31818cb278.

James, N. W., Adams, G. M. & Wilson, A. F. (1989). Determination of anaerobic threshold by ventilatory frequency. International journal of sports medicine 10 (3), 192–196. doi:10.1055/s-2007-1024899.

Kannankeril, P. J., Le, F. K., Kadish, A. H. & Goldberger, J. J. (2004). Parasympathetic effects on heart rate recovery after exercise. Journal of investigative medicine: the official publication of the American Federation for Clinical Research 52 (6), 394–401. doi:10.1136/jim-52-06-34.

Karapetian, G. K., Engels, H. J. & Gretebeck, R. J. (2008). Use of heart rate variability to estimate LT and VT. International journal of sports medicine, 29 (8), 652–657. doi:10.1055/s-2007-989423.

Karmali, S. N., Sciusco, A., May, S. M. & Ackland, G. L. (2017). Heart rate variability in critical care medicine: a systematic review. Intensive care medicine experimental, 5 (1), 33. doi:10.1186/s40635-017-0146-1.

Kenney, W. L., Willmore, J. H. & Costill, D. L. (2015) Physiology of Sports and Exercise. Sixth edition. Champaign: Human Kinetics.

Koo, T. K. & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. Journal of chiropractic medicine 15 (2), 155–163. doi:10.1016/j.jcm.2016.02.012.

MacRae, H. S., Dennis, S. C., Bosch, A. N. & Noakes, T. D. (1992). Effects of training on lactate production and removal during progressive exercise in humans. Journal of applied physiology, 72 (5), 1649–1656. doi:10.1152/jappl.1992.72.5.1649.

Martinmäki, K., Rusko, H., Saalasti, S. & Kettunen, J. (2006). Ability of short-time Fourier transform method to detect transient changes in vagal effects on hearts: a pharmacological blocking study. American journal of physiology. Heart and circulatory physiology 290 (6), H2582–H2589. doi:10.1152/ajpheart.00058.2005.

Massaro, S. & Pecchia, L. (2016). Heart Rate Variability (HRV) Analysis: A Methodology for Organizational Neuroscience. Organizational Research Methods 22. doi:10.1177/1094428116681072.

Mateo-March, M., Moya-Ramón, M., Javaloyes, A., Sánchez-Muñoz, C. & Clemente-Suárez, V. J. (2022). Validity of detrended fluctuation analysis of heart rate variability to determine intensity thresholds in elite cyclists. European journal of sport science 27, 1–8. doi:10.1080/17461391.2022.2047228.

Mendia-Iztueta, I., Monahan, K., Kyröläinen, H. & Hynynen, E. (2016). Assessment of Heart Rate Variability Thresholds from Incremental Treadmill Tests in Five Cross-Country Skiing Techniques. PloS one 11 (1), e0145875. doi:10.1371/journal.pone.0145875.

Michael, S., Jay, O., Halaki, M., Graham, K. & Davis, G. M. (2016). Submaximal exercise intensity modulates acute post-exercise heart rate variability. European journal of applied physiology 116 (4), 697–706. doi:10.1007/s00421-016-3327-9.

Mourot, L., Bouhaddi, M., Perrey, S., Rouillon, J. D. & Regnard, J. (2004). Quantitative Poincaré plot analysis of heart rate variability: effect of endurance training. European journal of applied physiology 91 (1), 79–87. doi:10.1007/s00421-003-0917-0.

Mourot, L., Fabre, N., Savoldelli, A. & Schena, F. (2014). Second ventilatory threshold from heart-rate variability: valid when the upper body is involved? International journal of sports physiology and performance 9 (4), 695–701. doi:10.1123/ijspp.2013-0286.

Myers, J. & Ashley, E. (1997). Dangerous curves. A perspective on exercise, lactate, and the anaerobic threshold. Chest 111 (3), 787–795. doi:10.1378/chest.111.3.787.

Nascimento, E., Antunes, D., do Nascimento Salvador, P. C., Borszcz, F. K. & de Lucas, R. D. (2019). Applicability of Dmax Method on Heart Rate Variability to Estimate the Lactate Thresholds in Male Runners. Journal of sports medicine 2019, 2075371. doi:10.1155/2019/2075371.

Nascimento, E., Kiss, M., Santos, T., Lambert, M. & Pires, F. (2017). Determination of Lactate Thresholds in Maximal Running Test by Heart Rate Variability Data Set. Asian Journal of Sports Medicine. In Press. doi:10.5812/asjsm.58480.

Novelli, F. I., de Araújo, J. A., Tolazzi, G. J., Tricot, G. K., Arsa, G. & Cambri, L. T. (2019). Reproducibility of Heart Rate Variability Threshold in Untrained Individuals. International journal of sports medicine 40 (2), 95–99. doi:10.1055/a-0800-8633.

Nummela, A. & Peltonen, J. (2018). Suorat testit. In a book edited by K. Keskinen et al. Fyysisen kunnon testaaminen – käsi- ja oppikirja kuntotestaajille. Helsinki: Liikuntatieteellinen Seura ry, 79–101.

Nummela, A. (2021). Sykerajat ja kynnysarvot kestävyysharjoittelussa. Referred 23.5.2023. https://sportslab.fi/sykerajat-ja-kynnysarvot-kestavyysharjoittelussa/.

Ohira, Y., Edgerton, V. R., Gardner, G. W., Gunawardena, K. A., Senewiratne, B. & Ikawa, S. (1981). Work capacity after iron treatment as a function of hemoglobin and iron deficiency. Journal of nutritional science and vitaminology 27 (2), 87–96. doi:10.3177/jnsv.27.87.

Parak, J. & Korhonen, I. (s.a.). Accuracy of Firstbeat Bodyguard 2 beat-to-beat heart rate monitor. Firstbeat White Paper. www.firstbeat.com/wp-content/uploads/2015/10/white_paper_bodyguard2_final.pdf.

Pfitzinger, P. & Freedson, P. S. (1998). The reliability of lactate measurements during exercise. International journal of sports medicine 19 (5), 349–357. doi:10.1055/s-2007-971929.

Peltonen, J. & Nummela, A. (2018). Kestävyyden fysiologiset perusteet. In a book submitted by K. Keskinen et al. Fyysisen kunnon testaaminen – käsi- ja oppikirja kuntotestaajille. Helsinki: Liikuntatieteellinen Seura ry, 64–78.

Queiroz, M.G., Arsa, G., Rezende, D.A., Sousa, L.C.J.L., Oliveira, F.R., Araujo, G.G. & Cambri, L.T. (2018). Heart rate variability estimates ventilatory threshold regardless

body mass index in young people. Science & Sports 33 (1), 39–46, doi:10.1016/j.scispo.2017.06.005.

Quinart, S., Mourot, L., Nègre, V., Simon-Rigaud, M. L., Nicolet-Guénat, M., Bertrand, A. M., Meneveau, N. & Mougin, F. (2014). Ventilatory thresholds determined from HRV: comparison of 2 methods in obese adolescents. International journal of sports medicine 35 (3), 203–208. doi:10.1055/s-0033-1345172.

Quirion, A., Brisson, G.R., Laurencelle, L., DeCarufel, D., Audet, A., Dulac, S., Ledoux, M. & Vogelaere, P. (1988) Lactate threshold and onset of blood lactate accumulation during incremental exercise after dietary modifications. European Journal of Applied Physiology 57, 192–197. doi:10.1007/BF00640662

Reinhard, U., Müller, P. H. & Schmülling, R. M. (1979). Determination of anaerobic threshold by the ventilation equivalent in normal individuals. Respiration; international review of thoracic diseases 38 (1), 36–42. doi:10.1159/000194056.

Robergs, R. A., Ghiasvand, F. & Parker, D. (2004). Biochemistry of exercise-induced metabolic acidosis. American journal of physiology. Regulatory, integrative and comparative physiology 287 (3), R502–R516. doi:10.1152/ajpregu.00114.2004.

Rodríguez-Marroyo, J. A., Villa, J. G., García-López, J. & Foster, C. (2013). Relationship between the talk test and ventilatory thresholds in well-trained cyclists. Journal of strength and conditioning research 27 (7), 1942–1949. doi:10.1519/JSC.0b013e3182736af3.

Rogers, B., Berk, S. & Gronwald, T. (2022). An Index of Non-Linear HRV as a Proxy of the Aerobic Threshold Based on Blood Lactate Concentration in Elite Triathletes. Sports (Basel, Switzerland) 10 (2), 25. doi:10.3390/sports10020025.

Rogers, B., Giles, D., Draper, N., Hoos, O. & Gronwald, T. (2021a). A New Detection Method Defining the Aerobic Threshold for Endurance Exercise and Training Prescription Based on Fractal Correlation Properties of Heart Rate Variability. Frontiers in physiology 11, 596567. doi:10.3389/fphys.2020.596567.

Rogers, B., Giles, D., Draper, N., Mourot, L. & Gronwald, T. (2021b). Detection of the Anaerobic Threshold in Endurance Sports: Validation of a New Method Using Correlation Properties of Heart Rate Variability. Journal of functional morphology and kinesiology 6 (2), 38. doi:10.3390/jfmk6020038.

Rogers, B., Mourot, L. & Gronwald, T. (2021c). Aerobic Threshold Identification in a Cardiac Disease Population Based on Correlation Properties of Heart Rate Variability. Journal of clinical medicine 10 (18), 4075. doi:10.3390/jcm10184075.

Rogers, B., Giles, D., Draper, N., Mourot, L. & Gronwald, T. (2021d). Influence of Artefact Correction and Recording Device Type on the Practical Application of a Non-Linear Heart Rate Variability Biomarker for Aerobic Threshold Determination. Sensors (Basel, Switzerland), 21 (3), 821. doi:10.3390/s21030821.

Ryan, S. M., Goldberger, A. L., Pincus, S. M., Mietus, J. & Lipsitz, L. A. (1994). Gender- and age-related differences in heart rate dynamics: are women more complex than men? Journal of the American College of Cardiology, 24 (7), 1700–1707. doi:10.1016/0735-1097(94)90177-5.

Sales, M. M., Campbell, C. S., Morais, P. K., Ernesto, C., Soares-Caldeira, L. F., Russo, P., Motta, D. F., Moreira, S. R., Nakamura, F. Y. & Simões, H. G. (2011). Noninvasive method to estimate anaerobic threshold in individuals with type 2 diabetes. Diabetology & metabolic syndrome 3 (1), 1. doi:10.1186/1758-5996-3-1.

Sandercock, G. R. & Brodie, D. A. (2006). The use of heart rate variability measures to assess autonomic control during exercise. Scandinavian journal of medicine & science in sports 16 (5), 302–313. doi:10.1111/j.1600-0838.2006.00556.x.

Shaffer, F., McCraty, R. & Zerr, C. L. (2014). A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability. Frontiers in psychology 5, 1040. doi:10.3389/fpsyg.2014.01040.

Schmidt, T., Wulff, S., Braumann, K.-M. & Reer, R. (2019). Determination of the Maximal Lactate Steady State by HRV in Overweight and Obese Subjects. Sports medicine international open 3, E58–E64. doi: 10.1055/a-0883-5473.

Sjödin, B. & Jacobs, I. (1981). Onset of blood lactate accumulation and marathon running performance. International journal of sports medicine, 2 (1), 23–26. doi:10.1055/s-2008-1034579.

Smekal, G., von Duvillard, S. P., Frigo, P., Tegelhofer, T., Pokan, R., Hofmann, P., Tschan, H., Baron, R., Wonisch, M., Renezeder, K. & Bachl, N. (2007). Menstrual cycle: no effect on exercise cardiorespiratory variables or blood lactate concentration. Medicine and science in sports and exercise 39 (7), 1098–1106. doi:10.1249/mss.0b013e31805371e7.

Stanley, J., Peake, J. M. & Buchheit, M. (2013). Cardiac parasympathetic reactivation following exercise: implications for training prescription. Sports medicine (Auckland, N.Z.) 43 (12), 1259–1277. doi:10.1007/s40279-013-0083-4.

Taipale-Mikkonen, R. S., Raitanen, A., Hackney, A. C., Solli, G. S., Valtonen, M., Peltonen, H., McGawley, K., Kyröläinen, H. & Ihalainen, J. K. (2021). Influence of Menstrual Cycle or Hormonal Contraceptive Phase on Physiological Variables Monitored During

Treadmill Testing. Frontiers in physiology 12, 761760. doi:10.3389/fphys.2021.761760.

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996). Heart Rate Variability: Standards of Measurement, Physiological Interpretation and Clinical Use. Circulation 93, 1043–1065.

Tulppo, M. P., Hughson, R. L., Mäkikallio, T. H., Airaksinen, K. E., Seppänen, T. & Huikuri, H. V. (2001). Effects of exercise and passive head-up tilt on fractal and complexity properties of heart rate dynamics. American journal of physiology. Heart and circulatory physiology 280 (3), H1081–H1087. doi:10.1152/ajpheart.2001.280.3.H1081.

Tulppo, M. P., Mäkikallio, T. H., Takala, T. E., Seppänen, T. & Huikuri, H. V. (1996). Quantitative beat-to-beat analysis of heart rate dynamics during exercise. The American journal of physiology 271 (1 pt 2), H244–H252. doi:10.1152/ajpheart.1996.271.1.H244.

Turnes, T., dos Santos, R. P., de Aguiar, R. A., Loch, T., Possamai, L. T. & Caputo, F. (2019). Association Between Deoxygenated Hemoglobin Breaking Point, Anaerobic Threshold, and Rowing Performance. International journal of sports physiology and performance, 14 (8), 1103–1109. doi:10.1123/ijspp.2018-0675.

Vanhatalo, A., Doust, J. H. & Burnley, M. (2007). Determination of critical power using a 3-min all-out cycling test. Medicine and science in sports and exercise, 39 (3), 548–555. doi:10.1249/mss.0b013e31802dd3e6.

Wasserman, K. & McIlroy, M. B. (1964). Detecting the threshold of anaerobic metabolism in cardiac patients during exercise. The American journal of cardiology, 14, 844–852. doi:10.1016/0002-9149(64)90012-8.

Wasserman, K., Van Kessel, A. L. & Burton, G. G. (1967). Interaction of physiological mechanisms during exercise. Journal of applied physiology 22 (1), 71–85. doi:10.1152/jappl.1967.22.1.71.

Wasserman, K., Whipp, B. J., Koyl, S. N. & Beaver, W. L. (1973). Anaerobic threshold and respiratory gas exchange during exercise. Journal of applied physiology 35 (2), 236–243. doi:10.1152/jappl.1973.35.2.236.

Yoshida, T. (1984) Effect of dietary modifications on lactate threshold and onset of blood lactate accumulation during incremental exercise. European Journal of Applied Physiology 53, 200–205. doi:10.1007/BF00776590.

Zimatore, G., Falcioni, L., Gallotta, M. C., Bonavolontà, V., Campanella, M., De Spirito, M., Guidetti, L. & Baldari, C. (2021). Recurrence quantification analysis of heart rate

variability to detect both ventilatory thresholds. PloS one 16 (10), e0249504. doi:10.1371/journal.pone.0249504.

APPENDIX 1: Background information form (translated)

UKK-instituutti
Tampereen Urheilulääkäriasema

terve urheilija

BACKGROUND SURVEY / RUN RCT

Name: _____

Date of birth: _____ Email: _____

1. ESTIMATED LEVEL OF PHYSICAL ACTIVITY

Pleas choose ONE NUMBER (0 – 7) which best depicts your physical activity level during the last month:

1 = 0 -1 h a week
2 = 1-3 h a week
3 = 4-7 h a week
4 = 8-10 h a week
5 = 10-15 h a week
6 = more than 15 h a week

Sports which you have been doing:

_____

2. My state of health HAS BEEN / HAS NOT BEEN the same after filling in the background survey of the current study.

If not, specify which kind of changes have occurred: _____

4. Have you felt crushing chest pain? No Yes

Has it worsened during physical activity? No Yes

7. Do you smoke or use snus regularly? No Yes

specify the amount of use: _____

8. Do you currently use some medications? No Yes

specify the names of the medications: _____

_____

9. Are you feeling extraordinally tired today? No Yes

If yes, why? _____

UKK-instituutti
Tampereen Urheilulääkäriasema

terve urheilija

10. Are you feeling yourself completely healthy? Yes No

If no, please describe your symptoms:_____

11. If you did training(s) during the preceding two days, please describe what did you do

yesterday _____

the day before yesterday _____

12. Time after the latest portition of alcohol _____ h

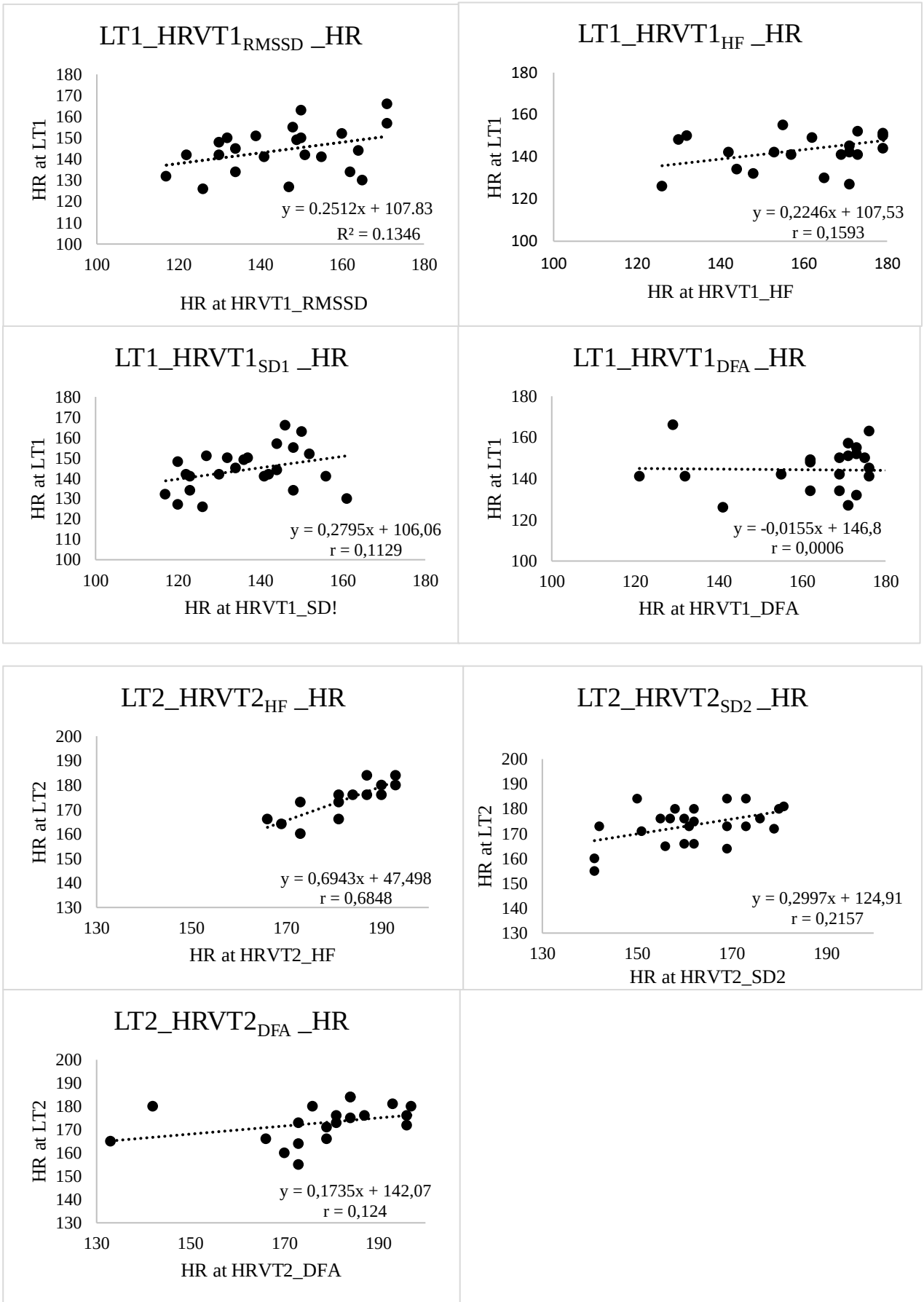
13. Time after the latest portition of caffeinated beverage (coffee etc.)_____ h

WE ARE STORING THE PERSONAL AND TEST DATA FOR 10 YEARS IN OUR CUSTOMER AND TESTING DATA REGISTER.

WITH MY SIGNATURE, I CONFIRM THAT I HAVE GIVEN TRUTHFUL INFORMATION ABOUT MY STATE OF HEALTH.

Place and date_____ **Signature**_____

APPENDIX 2: Linear regression graphs of HR comparisons between LTs and HRVTs.



APPENDIX 3: Linear regression graphs of speed comparisons between LTs and HRVTs.