



## OPEN ACCESS

## EDITED BY

Emiliano Cè,  
University of Milan, Italy

## REVIEWED BY

Jean-Marc Vesin,  
Swiss Federal Institute of Technology  
Lausanne, Switzerland  
Raquel Bailón,  
University of Zaragoza, Spain

## \*CORRESPONDENCE

Matias Kanniainen,  
✉ matias.kanniainen@tuni.fi

RECEIVED 22 September 2023

ACCEPTED 30 November 2023

PUBLISHED 21 December 2023

## CITATION

Kanniainen M, Pukkila T, Kuisma J,  
Molkkari M, Lajunen K and Räsänen E  
(2023), Estimation of physiological  
exercise thresholds based on dynamical  
correlation properties of heart rate  
variability.  
*Front. Physiol.* 14:1299104.  
doi: 10.3389/fphys.2023.1299104

## COPYRIGHT

© 2023 Kanniainen, Pukkila, Kuisma,  
Molkkari, Lajunen and Räsänen. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which does  
not comply with these terms.

# Estimation of physiological exercise thresholds based on dynamical correlation properties of heart rate variability

Matias Kanniainen<sup>1\*</sup>, Teemu Pukkila<sup>1</sup>, Joonas Kuisma<sup>1</sup>,  
Matti Molkkari<sup>1</sup>, Kimmo Lajunen<sup>2</sup> and Esa Räsänen<sup>1</sup>

<sup>1</sup>Computational Physics Laboratory, Tampere University, Tampere, Finland, <sup>2</sup>Kauppi Sports Coaching Ltd., Tampere, Finland

Aerobic and anaerobic thresholds of the three-zone exercise model are often used to evaluate the exercise intensity and optimize the training load. Conventionally, these thresholds are derived from the respiratory gas exchange or blood lactate concentration measurements. Here, we introduce and validate a computational method based on the RR interval (RRI) dynamics of the heart rate (HR) measurement, which enables a simple, yet reasonably accurate estimation of both metabolic thresholds. The method utilizes a newly developed dynamical detrended fluctuation analysis (DDFA) to assess the real-time changes in the dynamical correlations of the RR intervals during exercise. The training intensity is shown to be in direct correspondence with the time- and scale-dependent changes in the DDFA scaling exponent. These changes are further used in the definition of an individual measure to estimate the aerobic and anaerobic threshold. The results for 15 volunteers who participated in a cyclo-ergometer test are compared to the benchmark lactate thresholds, as well as to the ventilatory thresholds and alternative HR-based estimates based on the maximal HR and the conventional detrended fluctuation analysis (DFA). Our method provides the best overall agreement with the lactate thresholds and provides a promising, cost-effective alternative to conventional protocols, which could be easily integrated in wearable devices. However, detailed statistical analysis reveals the particular strengths and weaknesses of each method with respect to the agreement and consistency with the thresholds—thus underlining the need for further studies with more data.

## KEYWORDS

heart rate variability, time series analysis, detrended fluctuation analysis, aerobic threshold, anaerobic threshold, wearable health technology

## 1 Introduction

Increasing availability and popularity of health technology including wearable devices such as wrist monitors, rings and smart clothing brings significant possibilities to analyze the physiological signals in everyday life. There are multiple different measures to evaluate the effects of physical exercise from the collected data alone. One of the applications from the exercise data is to optimize the training load, and determine the different physiological zones during the exercise. Currently, there are several measures to determine different physiological changes during exercise including the conventional measurements of oxygen

consumption (Foster, 1983) and lactate concentration of the blood (Faude et al., 2009), as well as estimates based on the heart rate (HR) and HR variability (HRV) measurements (Cottin et al., 2006; van Campen et al., 2020) with varying reliability and usefulness.

In a three-zone model of the exercise (Stöggl and Sperlich, 2019), the training zones are separated by thresholds, which can be either determined through the ventilatory exchange ( $VT_1$ ,  $VT_2$ ) (Beaver et al., 1986), or the lactate concentration ( $LT_1$ ,  $LT_2$ ) (Binder et al., 2008). The first ( $TH_1$ ) and second thresholds ( $TH_2$ ) are known as the aerobic and anaerobic thresholds, respectively. There are multiple different ways to determine the lactate thresholds from the lactate concentration (Newell et al., 2007; Jamnick et al., 2018; Jamnick et al., 2020), which brings uncertainty in the analysis. Furthermore, these invasive measurements are costly and time-consuming. On the other hand, the non-invasive measurements of the ventilatory exchange require a strict set of quality-control criteria (Gaskill et al., 2001; Cannon et al., 2009) and furthermore requires a specialized test environment with trained personnel to perform the measurements during the exercise.

Most wearable devices use the relative HR compared to the estimated maximal HR to determine the training zones. With this method the aerobic threshold is usually estimated to be around 60%–70% of the maximal HR ( $HR_{max}$ ) and the anaerobic threshold around 85%–90% of the  $HR_{max}$ , respectively. However, this estimation has a few shortcomings. First, the  $HR_{max}$  has significant individual variability, so the simple age-dependent models are not accurate or universal (Shookster et al., 2020). Secondly, even when using the actual measured individual  $HR_{max}$  the percentage of the  $HR_{max}$  is also highly individual. So this model only works for population averages but fails to accurately detect individual thresholds. This opens a huge need and market for more accurate detection methods.

To date, there is no golden standard in the determination of the metabolic thresholds. There is variation in the preference, depending also on the exercise protocols and their availability on the market. The Finnish standard for threshold determination is based on the lactate values (Keskinen et al., 2018), and single case studies (Jones, 1998; Tjelta, 2019) support the use of the lactate concentration to determine the thresholds. However, there are publications and standards (Hoffman, 1999; Blain et al., 2005) which use the ventilatory thresholds as a baseline measure for the thresholds. Often, these conventional methods give quantitatively different results (Yeh et al., 1983; Şekir et al., 2002), and they are also dependent on subjective visual analysis included in the interpretation of the results, e.g., in the fitting procedures of linear trends. At present, there are new computerized methods available to define the training zones (Kim et al., 2021; Zignoli, 2023), but a simple, cost-effective and accurate method is still to be found.

HRV is a physiological measure, which captures the variation of the cardiac interbeat intervals (IBI) (Goldberger et al., 2002). The physiological responses of the exercise during both rest and exercise can be analyzed with HRV (Gronwald and Hoos, 2020). There are many different HRV metrics, often divided into time-domain, frequency-domain and non-linear measures (Shaffer and Ginsberg, 2017). There are several studies examining the determination of

the physiological thresholds using the frequency-domain methods (Di Michele et al., 2012; Ramos-Campo et al., 2017). Regarding non-linear HRV methods, it was recently suggested that the short-term scaling exponent  $\alpha_1$  of detrended fluctuation analysis (Peng et al., 1995) (DFA), i.e., a measure for the characteristics of the RR interval (RRI) correlations for the time scale of 4–16 consecutive RR intervals, provides a simple yet accurate approach to determine the ventilation thresholds (Rogers et al., 2021a; Rogers et al., 2021b).

In this study, we utilize a modified version of the DFA, i.e., dynamical detrended fluctuation analysis (Molkkari et al., 2020) (DDFA) to derive an approximation for both aerobic and anaerobic thresholds. DDFA allows the monitoring of the scaling exponent and thus the characteristics of the RR interval correlations as functions of both scale  $s$  and time  $t$ . In particular, it was shown by some of the present authors that under running  $\alpha(t, s)$  gradually decreases from small up to higher scales, and high-intensity exercise can be characterized by anticorrelations of the RR intervals, especially for small scales (Molkkari et al., 2020). The DDFA results can be further sorted according to the HR to examine  $\alpha(HR, s)$ . Here, this quantity was used to derive a simple yet effective measure for the aerobic and anaerobic threshold. The method was tested with HR data obtained from 15 subjects during a cyclo-ergometer test and compared to simultaneous lactate and ventilation measurements. Overall, our method yields the smallest overall difference from the LTs, thus overperforming other HR-based estimations. The VTs, instead, give fairly consistent results, e.g., relatively small variations across the samples with, however, systematic and significant overestimation of both LTs.

Our further statistical analysis shows that the introduced method is relatively consistent over the HR range of the subject-specific thresholds, even though the number of subjects (15) limits the conclusions. In summary, our method provides a promising tool to determine the physiological zones and the corresponding thresholds during training. We further discuss the potential and applicability of the method in wearable technologies.

## 2 Materials and methods

### 2.1 Participants

The participants of the study are 15 healthy volunteers (aged 22–44), who performed an exercise test at Kauppi Sports Coaching Ltd. The participants filled a form stating their legal gender, age, medical risk factors, exercise background and training goals. The individual information and maximal HR,  $VO_2$ , power and the RR interval filtering (described below) percentages are shown in Table 1. The participants were instructed to only eat lightly and not to consume caffeine for a few hours or alcohol for a few days before coming to the test. The exercise backgrounds of the subjects were taken into account when choosing the power levels for the test. The participants gave a written consent for the study. The approval of the study was given by the Tampere University Hospital Ethics Committee, and the principles of the Declaration of Helsinki were followed.

**TABLE 1** Subject-specific information including the gender (male = M, female = F), age (years), maximal heart rate:  $HR_{max}$  (BPM), maximal VO2 intake:  $VO2_{max}$  (mL/min), maximal ergometer power:  $Power_{max}$  (W), exercise duration (min:sec) and filtered RR intervals (%) for the studied subjects.

Subject	Gender	Age	$HR_{max}$	$VO2_{max}$	$Power_{max}$	Duration	Filtered RRI
1	M	33	195	3715	275	26:29	15.0
2	M	36	193	4353	370	27:19	0.2
3	M	44	204	3998	370	26:48	0.6
4	M	37	182	4139	370	29:05	0.0
5	F	27	202	2611	240	27:00	0.0
6	M	31	200	3911	300	29:11	0.0
7	F	35	172	2642	200	24:06	0.1
8	F	41	185	2687	200	21:33	0.0
9	F	22	200	3171	220	25:43	0.0
10	F	34	186	3152	250	24:31	0.1
11	F	30	185	3276	280	30:31	0.1
12	F	35	172	3224	225	19:54	5.6
13	F	42	176	2650	220	25:47	0.2
14	M	42	184	4132	330	23:52	0.1
15	M	27	202	4639	390	27:00	0.0
mean	M = 7/F = 8	34	189	3487	283	25:55	1.5
( $\pm$ std)		( $\pm$ 6)	( $\pm$ 11)	( $\pm$ 611)	( $\pm$ 66)	( $\pm$ 02:43)	( $\pm$ 3.9)

## 2.2 Test protocol and RR interval measurement

Before the exercise test weight, height and body composition of the participants were measured. The test was performed with a Monark LC4 cyclo-ergometer, and the test started after a 5-min warm up period. Based on the training background of the participant, the test was commenced with a power of 40–120 W and every 3 min the power was increased incrementally by 20–30 W. Individually chosen power levels were chosen to best suit each participant's fitness level. The choice is based on Keskinen et al. (2018), where the change of one level is derived from the body weight of the subject. As suggested by American College of Sports Medicine (2013) et al., the maximal power level is estimated as the eight increase of the power level, and the starting level is therefore determined with (Starting level = Estimated level(W) – 7 × change of one level(W)). The test was done until exhaustion, when the subjects could not keep the required power level. During the measurement, the RR intervals were measured with a Polar H10 heart rate sensor, which has been shown to be highly reliable for RRI determination, especially during intense physical exercise (Gilgen-Ammann et al., 2019; Schaffarczyk et al., 2022). We considered the RR intervals

only during the exercise, so the analysis does not cover the warm-up period. Polar H10 was connected to the Polar M430 watch for data storage. After the test, the RRI data was transferred to the computer for analysis, which was performed with Python.

## 2.3 Preprocessing of RR intervals

First, we ignored the RRI values outside of the data range of 200–2,000 ms, or correspondingly 30–300 BPM in the HR. Then automatic filtering was performed by ignoring values that deviated by more than 10% from windowed median filter with kernel size of 7 RRIs. Data filtering in HRV measurements is a commonly known problem (dos Santos et al., 2013) and it is challenging to develop a robust and universal filtering algorithm. We did not have the access to the full ECG recordings and there is no certainty about the nature of the outliers. Therefore, we visually examined the samples and removed the beats which were statistically outliers in the current context not to cause problems in the following analysis. A few apparently incorrect missing beats or extra R peaks not detected by the automatic algorithm were removed by visual inspection.

The filtered percentages of the RRI data are shown in Table 1. Generally, the data was of excellent quality and little to none

removal of the intervals was required for the majority of the samples. However, a significant amount (15.0%) of the RRI of subject 1 were removed. Despite this removal, the filtered data of subject 1 was also used in the analysis with relatively little effect on the results.

## 2.4 Lactate and ventilation measurements

To measure blood lactate concentrations, fingertip blood samples were taken at the end of each 3-min exercise load. These samples were analyzed with EKF Biosen C-line Clinic device. In the definition of the thresholds through lactate measurements we followed the standard protocol used in Finland, refined over the years (Keskinen et al., 2018). The data points from the lactate measurements (mmol/L) are plotted as a function of time and visual analysis is performed.  $LT_1$  is defined as 0.3 mmol/L above the lowest lactate level (by linear interpolation between two lactate measurements, if necessary). Then the remaining data points are split into two linear fits: the first fit is made between  $LT_1$  and the following data point, and the second fit is made through the final points where the lactate concentration increases by more than 0.8 mmol/L during the exercise step.  $LT_2$  is defined as the crossing point of the two slopes (Keskinen et al., 2018).

The ventilatory exchange was measured with Cosmed Quark CPET respiratory gas analyzer during the exercise. The ventilatory thresholds were defined by combining multiple automatic methods and taking the average, as described in Kim et al. (2021).  $VT_1$  is the average of the V-slope and excess carbon dioxide methods, whereas  $VT_2$  is the average of the V-slope and excess minute ventilation methods (Kim et al., 2021).

Furthermore, we converted the thresholds from time to HR by taking a corresponding HR value from the measurement results. During the measurement, the HR was averaged with a 30-s window, so the HR distribution was stable and thus easily convertible.

## 2.5 Thresholds from heart rate and heart rate variability

### 2.5.1 Maximal HR method

We calculated the thresholds based on maximal HR analysis, where the first threshold ( $HR_{maxT_1}$ ) was set at 70% of the  $HR_{max}$ , and the second threshold ( $HR_{maxT_2}$ ) was set at 85% of the  $HR_{max}$  (McPartland et al., 2010). These values are often used by different manufacturers of the wearable devices, in which physiological threshold estimates are built in. However, some manufacturers use different values and even determine the exercise zones in a different way compared to that of Stöggl and Sperlich (2019). Nevertheless, we resort to the most common values given above, but point out that there is significant individual variation in the percentage of  $HR_{max}$  for both thresholds depending on, e.g., the training background and age. This is a subject of assessment also in this work (see Section 3.2).

### 2.5.2 DFA $\alpha_1$ method

We compared our threshold estimation method with that of Rogers et al. (2021a); Rogers et al. (2021b) by calculating the conventional DFA  $\alpha_1$  every 5 seconds in 2-min segments as a function of the average HR within the segments. Smoothness priors

detrending (Tarvainen et al., 2002) was applied to the RRI time series (smoothing parameter  $\lambda = 500$ ), but other preprocessing was conducted according to Section 2.3 instead of the Kubios software as described in Rogers et al. (2021a); Rogers et al. (2021b). The algorithm relied on linear regression on the HR versus  $\alpha_1$  in the region of rapid near-linear decline of  $\alpha_1$  to determine the aerobic and anaerobic thresholds wherein the regression line crossed  $\alpha_1 = 0.75$  and  $\alpha_1 = 0.5$ , respectively. However, the definition of the regression region was left subjective in Rogers et al. (2021a); Rogers et al. (2021b). Therefore, we adopted the following algorithmic approach to support systematic studies:

- 1) Initially select all the data points that fall within the interval  $0.5 \leq \alpha_1(\text{HR}) \leq 0.75$ .
- 2) If there are multiple disjoint regions (as a function of HR) where this condition is fulfilled, connect the regions by also selecting the intervening data points if there are at most  $n$  of them. Here we chose  $n = 4$ .
- 3) If there are still multiple disjoint regions, choose the one with the most data points. Perform linear regression over this region and record the coefficient of determination.
- 4) Expand the region one data point at a time from either end, and choose the one region that results in the largest coefficient of determination for the linear regression.

The performance of the algorithm was visually inspected to ascertain that it produced sensible results. We point out that this method was originally developed to assess the VTs (instead of LTs). This is taken into account in the interpretation of the results below.

### 2.5.3 Dynamical DFA method

Our computational method introduced below to estimate the metabolic thresholds from the RR intervals is based on DDFA Molkkari et al. (2020) with recent improvements (Molkkari and Räsänen, 2023). We utilize the second-order DDFA, where a second-order polynomial is fitted in the conventional DFA fluctuation function instead of a linear fit. The DDFA scaling exponents are calculated as follows.

- 1) Perform dynamic segmentation for each scale  $s$ , where the segment length  $l = 5s$ .
- 2) Compute the second-order DFA fluctuation function in maximally overlapping windows for each segment at scales  $s - 1, s, s + 1$ . The fluctuation function is thus denoted as  $\tilde{F}_t(s - 1)$ ,  $\tilde{F}_t(s)$  and  $\tilde{F}_t(s + 1)$ , for the respective scales.
- 3) In each segment, compute the dynamic scaling exponent  $\alpha(t, s)$  by the finite difference approximation

$$\alpha(t, s) \approx \frac{[h_-^2 \tilde{F}_t(s + 1) + (h_+^2 - h_-^2) \tilde{F}_t(s) - h_+^2 \tilde{F}_t(s - 1)]}{[h_- h_+ (h_+ + h_-)]},$$

where  $h_- = \log(s) - \log(s - 1)$  and  $h_+ = \log(s + 1) - \log(s)$  are the logarithmic backward and forward differences.

The metabolic thresholds are derived from the resulting DDFA scaling exponents, where the scaling exponents are aggregated as a function of HR. The procedure consists of the following steps.

- 1) Calculate the second-order DDFA (DDFA-2) for the RRI time series with 20 logarithmically spaced integer scales between 5–64 RRIs. The range of scales corresponds to the to the joint scales of DFA  $\alpha_1$  and  $\alpha_2$ .

- 2) Calculate the scaling exponents  $\alpha(\text{HR}, s)$  as a function of HR and scale  $s$ , where HR is the average of each segment. Sort the HR values by binning them to the nearest integer HR and assign  $\alpha(\text{HR}, s)$  values to the corresponding bins. For each scale  $s$ , take the mean value of the scaling exponents within the bins yielding a distribution of  $\alpha(\text{HR}_{\text{bin}}, s)$ .
- 3) Calculate an individual baseline to reduce the variability of different physiological starting conditions. The baseline is determined from the mean value of  $\alpha(\text{HR}_{\text{bin}}, s)$  for the 25 bins with the lowest HR for each scale  $s$ . The baseline values of each scale are subtracted from  $\alpha(\text{HR}_{\text{bin}}, s)$  over the whole measurement.
- 4) Calculate the mean value of  $\alpha(\text{HR}_{\text{bin}}, s)$  over each scale  $s$  of a HR bin. Smoothen the resulting  $\alpha(\text{HR}_{\text{bin}})$  curve with a mean filter with a kernel size of 10 HR bins to prefer the trend over small local fluctuations obtaining the mean smoothed scaling exponent  $\bar{\alpha}(\text{HR}_{\text{bin}})$ .

The first DDEFA derived threshold (DDFAT<sub>1</sub>) is the point, where  $\bar{\alpha}(\text{HR}_{\text{bin}})$  distribution drops below the baseline. The determination of this point is derived through the following calculation. First find the points where  $\bar{\alpha}(\text{HR}_{\text{bin}})$  distribution crosses the baseline. If the intersection is not stable, i.e.,  $\bar{\alpha}(\text{HR}_{\text{bin}})$  keeps fluctuating around the baseline, move to the next one until a stable intersection is found. We find that sufficient stability can be found when  $\bar{\alpha}(\text{HR}_{\text{bin}})$  remains negative for at least 10 consecutive binned HR values.

- 5) Similarly, select the second DDEFA derived threshold (DDFAT<sub>2</sub>) in a stable intersection where  $\bar{\alpha}(\text{HR}_{\text{bin}})$  equals  $-0.5$ .

Thus,  $\bar{\alpha}(\text{HR}_{\text{bin}})$  measures physiological changes compared to the baseline scaling exponent, according to the previously observed decrease in the scaling exponent under physical exercise (Karasik et al., 2002; Hautala et al., 2003; Gronwald et al., 2019; Molkkari et al., 2020).

## 2.6 Statistical analysis

We computed Pearson correlation coefficients for the different threshold determination methods against the LTs. As the correlations were expected to be positive, we considered the corresponding one-sided alternative hypothesis against the null hypothesis of no correlation. Because of the small sample size, we computed the  $p$ -values with a random permutation test.

In addition to correlations, we assessed *agreement* between the threshold estimates and LTs by performing an analysis akin to Bland and Altman (1999). Assuming normality of the differences, the 95% limits of agreement (LoA) were computed as the mean difference  $\pm 1.96$  standard deviations.

The uniformity of the differences throughout the range of measurements was assessed by linear regression. The test statistic was the absolute value of the regression slope, which was compared against the regression slopes of Monte Carlo-sampled uncorrelated normally distributed differences with the observed standard deviation. We calculated the  $p$ -values for the null hypothesis that the absolute value of the slope is drawn from normally distributed uncorrelated differences with the same standard deviation as the real differences *versus* the one-sided alternative hypothesis that the

absolute value of the slope is greater than what is expected for those normally distributed uncorrelated differences. Furthermore, we calculated the  $p$ -values for the Shapiro-Wilk normality test, where again the null hypothesis is that the differences are drawn from the normal distribution.

For all the statistics, 95% confidence intervals were computed with bias-corrected and accelerated bootstrapping (Efron, 1987). In all the methods that relied on random (re)sampling, the sampling was performed  $10^4$  times.

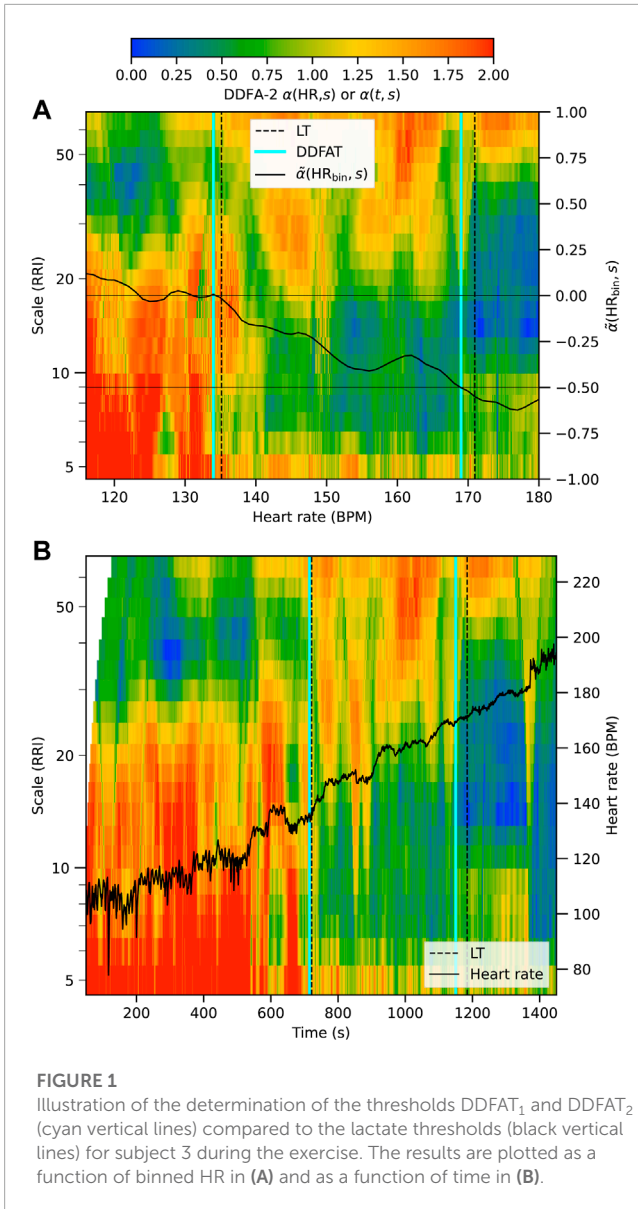
## 3 Results

### 3.1 Demonstration of threshold estimation

As described above, our method for the threshold estimation utilizes the dynamical RRI correlations through the DDEFA scaling exponent  $\alpha(\text{HR}, s)$ . Figure 1 illustrates our method for subject 3 during the exercise. The colors in the figure correspond to the values of the second-order DDEFA scaling exponent. In Figure 1A the scaling exponent is shown as functions of HR ( $x$ -axis) and scale ( $y$ -axis), whereas in Figure 1B it is given as functions of time and scale. The transition from higher to lower  $\alpha$  values during the exercise, or with increasing HR, is evident in colors that turn from red to blue across an increasing range of scales. The solid line in Figure 1A shows the mean  $\bar{\alpha}(\text{HR}_{\text{bin}})$  used in the determination of DDFAT<sub>1</sub> and DDFAT<sub>2</sub> as described above. The corresponding estimates for the thresholds are shown as cyan vertical dashed lines, and they can be compared with the benchmark LT values shown in black vertical dashed lines. In this particular example DDFAT thresholds underestimate LT<sub>1</sub> and LT<sub>2</sub> by only one and three BPM, respectively.

Next we consider the behavior of the RRIs during the exercise over *all* the subjects. These results can be qualitatively compared with the distribution of the LT values computed from the lactate measurements. Figure 2 shows an aggregate plot of  $\alpha(\text{HR}, s)$  for all the subjects (upper panel) together with the LTs (lower panel) as a function of HR. At low HR—corresponding here to the beginning of the exercise test—the scaling exponent value is relatively high. The behavior is in accordance with the DFA results previously obtained for a healthy heart at rest Goldberger et al. (2002). At higher HR the scaling exponent decreases, first at scales of 10–20 RRIs, and then through an extending range of scales with increasing HR. At high HR above about 160 BPM, the scaling exponent decreases well below 0.5, indicating anticorrelated behavior in line with the previous findings (Molkkari et al., 2020).

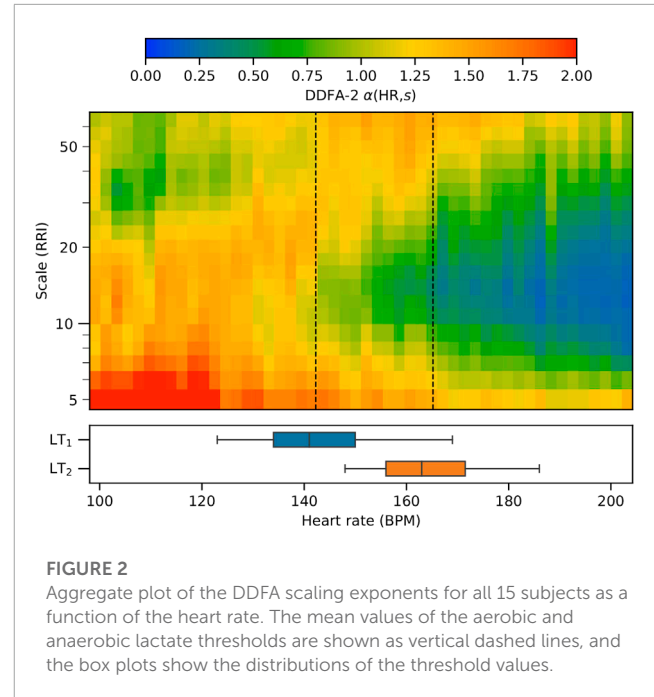
On the average, LT<sub>1</sub> qualitatively corresponds with the point where the scaling exponent begins to decrease below the baseline around 10 RRIs. On the other hand, LT<sub>2</sub> is approximately located in the area where the scaling exponent decreases below 0.5 from the baseline. The value is selected as the mean scaling exponents over the scales at rest corresponds to a value of around 1.0, thus a reduction of 0.5 corresponds to an arise of anticorrelated behaviour in the RRI correlations, which is associated with fatigue (Molkkari et al., 2020). This average behavior is in accordance with the threshold determination method described in the previous section. We point out, however, that as there is a large individual variance in the HRs corresponding to the respective LTs as demonstrated in the box plots of Figure 2. Quantitative comparison between the individuals,



including different threshold estimation methods, is presented in the following section.

### 3.2 Comparison of the methods

Table 2 shows the subject-specific HRs for both the first (aerobic) and second (anaerobic) thresholds computed with all the methods. The mean values and standard deviations are shown on the last two rows, respectively. As the first observation, the HRs of the LTs are significantly lower than those of VTs, the mean differences being 20 and 12 BPM for the first and second thresholds, respectively. This is in line with previous findings Yeh et al. (1983), although recent studies in treadmill tests have showed rather good agreement between LTs and VTs (Neves et al., 2022). Secondly, Table 2 allows the examination of different thresholds for the same subject, which is not visible on the visualizations. For example, for subject 5, LTs and DDFATs are close to each other and VTs and  $DFA\alpha_1$ Ts are



close to each other, but  $HR_{max}$ Ts are heavily underestimated. On the other hand, for subject 13 all the other thresholds are relatively close to each other, but once again VTs and  $DFA\alpha_1$ Ts have significantly higher values for both thresholds.

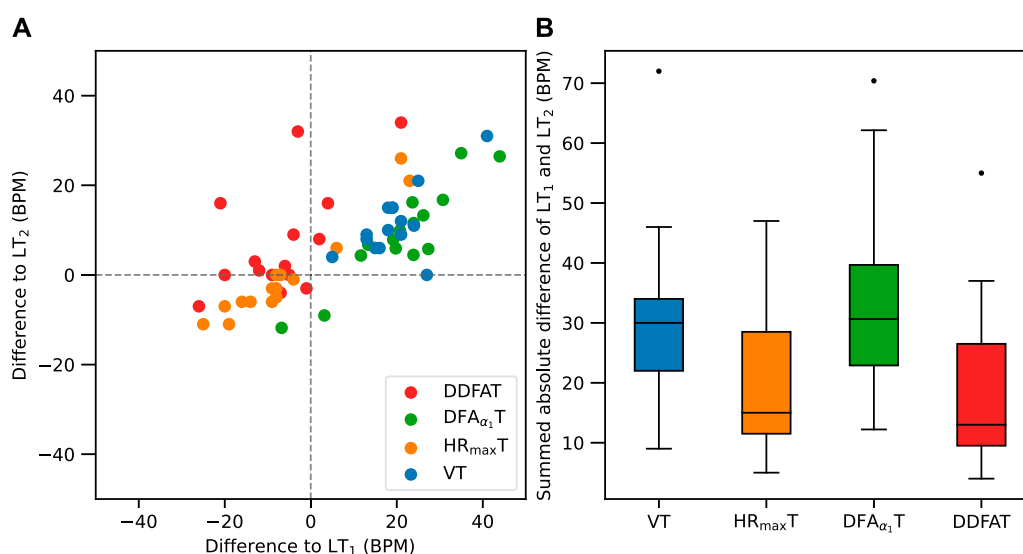
The differences of the methods from both  $LT_1$  and  $LT_2$  (in BPM) are visualized in Figure 3A. Most of the methods display relatively large variations with distinctive outliers for both thresholds. Some correlation between the differences with respect to  $LT_1$  and  $LT_2$ , especially for VT and  $DFA\alpha_1$ T is also visible. The summed absolute differences to illustrate the overall performance with respect to both LTs are shown in Figure 3B. The  $DFA\alpha_1$ T results agree well with VTs and thus systematically exceed the LT thresholds, leading to relatively high distances from the baseline in Figure 3B. On the other hand,  $HR_{max}$ T shows reasonable overall agreement, but include a few distinctive outliers. DDFAT shows relatively good agreement with the LT values with the smallest summed deviation in Figure 3B.

In addition to mere comparison between the HR estimates for the thresholds, it is worthwhile to examine the correlations between the estimates and the LTs and their statistical significance. Table 3 shows the Pearson correlation coefficients  $r$  with their bootstrapped 95% confidence interval lower bounds (LB) and the  $p$ -values for Pearson correlation against LTs ( $p_{corr}$ ). The VTs correlate the most with the LTs with the best statistical significance, whereas all the other methods show low to moderate correlation. We point out, however, that with the present data size correlation analysis is considerably affected by individual outliers present in the data.

Next, we utilize Bland–Altman plots to analyze the consistency in the comparisons between the methods in further detail. Figure 4 shows the differences from  $LT_1$  Figure 4A and  $LT_2$  Figure 4B as a function of HR with several statistical measures. In particular, the dark gray lines show the mean differences  $\mu_{diff}$  from the LTs listed also in Table 3, and the ends in the lines indicate their 95% confidence intervals calculated with bootstrapping. Correspondingly, the light grey lines show the 95% limits of

**TABLE 2** Subject-specific heart rates (in BPM) for the first (aerobic) and second (anaerobic) thresholds obtained by the different methods, along with the population mean and standard deviation.

Subject	LT <sub>1</sub>	VT <sub>1</sub>	HR <sub>max</sub> T <sub>1</sub>	DFA <sub>α<sub>1</sub></sub> T <sub>1</sub>	DDFAT <sub>1</sub>	LT <sub>2</sub>	VT <sub>2</sub>	HR <sub>max</sub> T <sub>2</sub>	DFA <sub>α<sub>1</sub></sub> T <sub>2</sub>	DDFAT <sub>2</sub>
1	150	171	141	153	138	177	186	171	168	178
2	145	160	137	157	132	171	177	166	175	174
3	135	162	158	162	134	171	171	192	177	168
4	144	157	130	165	146	163	172	157	173	171
5	169	182	144	188	163	186	194	175	194	188
6	123	164	144	158	144	148	179	174	175	182
7	131	149	122	155	126	151	161	148	163	151
8	125	150	131	169	122	153	174	159	179	185
9	150	169	142	181	154	172	187	172	189	188
10	153	158	133	146	127	169	173	162	157	162
11	140	164	132	166	119	163	174	160	176	179
12	141	157	122	161	132	159	165	148	165	159
13	133	151	126	157	129	153	168	153	169	162
14	135	156	131	159	128	161	173	160	165	157
15	160	179	144	173	140	181	196	175	188	181
mean	142	162	136	163	136	165	177	165	174	172
std	12.7	9.88	9.78	10.9	12.1	11.4	10.1	11.9	10.3	12.0



**FIGURE 3** (A) Comparison of differences between LT<sub>1</sub> (x-axis) and LT<sub>2</sub> (y-axis) for each subject with different threshold estimation methods. (B) Box plot of the summed absolute differences from both LT<sub>1</sub> and LT<sub>2</sub>.

**TABLE 3** Statistical data of different estimates against LT values: Pearson correlation coefficients  $r$ , their 95% confidence interval lower bounds (LBs),  $p$ -values for the null hypothesis of no correlation, mean difference  $\mu_{diff}$ , its 95% confidence interval LBs,  $p$ -values for the null hypothesis of zero slope ( $p_{slope}$ ) w.r.t. HR, and  $p$ -values for the null hypothesis that the differences are normally distributed with Shapiro–Wilk normality test ( $p_{norm}$ ).

Threshold	Pearson $r$	95% LB	$p_{corr}$	$\mu_{diff}$	$p_{slope}$	$p_{norm}$
VT <sub>1</sub>	0.78	0.37	$4.5 \times 10^{-4}$	19.7	0.15	0.18
HR <sub>max</sub> T <sub>1</sub>	0.28	-0.14	0.16	-6.5	0.30	0.031
DFA $\alpha_1$ T <sub>1</sub>	0.46	-0.17	0.046	21.1	0.51	0.72
DDFAT <sub>1</sub>	0.57	0.085	0.015	-6.7	0.81	0.60
VT <sub>2</sub>	0.76	0.35	$4.9 \times 10^{-4}$	11.5	0.48	0.18
HR <sub>max</sub> T <sub>2</sub>	0.58	0.13	0.012	5.8	0.83	0.0017
DFA $\alpha_1$ T <sub>2</sub>	0.52	0.034	0.025	9.0	0.64	0.38
DDFAT <sub>2</sub>	0.43	-0.035	0.053	7.1	0.83	0.014

agreement and their 95% confidence intervals. In addition, the figure shows the linear fits to data as a function of HR in colored lines together with their pointwise 95% confidence intervals shown in shaded areas. The statistical measure for the linearity is assessed by considering a null hypothesis of zero slope, and the corresponding  $p$ -values ( $p_{slope}$ ) are shown in Table 3. Here high  $p$ -values indicate a small degree of linearity. In that case the slope is consistent with normally distributed differences, implying a low chance of a trend. Table 3 shows also the  $p$ -values for the null hypothesis that the differences are normally distributed with Shapiro–Wilk normality test ( $p_{norm}$ ). The deviations from normality appear to be mainly due to few outliers in an already limited set of studied individuals, instead of skewness or other properties of the distributions. This underlines the fact that the analysis of the consistency of the results is cumbersome with the available amount of data.

Several observations can be made from Figure 4. First, the systematic overestimation of both LT<sub>1</sub> and LT<sub>2</sub> by VT<sub>1</sub> and VT<sub>2</sub>, respectively, is clearly visible, even though the correlation between LT and VT values is relatively high as noted above, i.e., VT provides consistent results in this respect. In addition, VT<sub>1</sub> shows the clearest linear trend with  $p_{slope} = 0.15$  in comparison with the other methods. HR<sub>max</sub>T<sub>1</sub> shows also a linear trend, although its uncertainty is large, mainly due to a few distinctive outliers. The overall performance of HR<sub>max</sub>T is relatively good, especially for LT<sub>2</sub>. However, the applicability of the estimate depends of the percentages of HR<sub>max</sub> selected for the thresholds (here 70% and 85%), and these percentages depend on individual fitness levels.

Figure 4 also shows that—as expected and noted also above—DFA $\alpha_1$ T values are generally in line with VTs but the mean difference from the LTs is relatively large, especially for the first threshold. In contrast with VTs, however, DFA $\alpha_1$ T values do not show linear bias. Finally, DDFAT shows reasonable agreement with LTs for both thresholds without linear bias; the values of  $p_{slope}$  are the highest in DDFAT of all the methods.

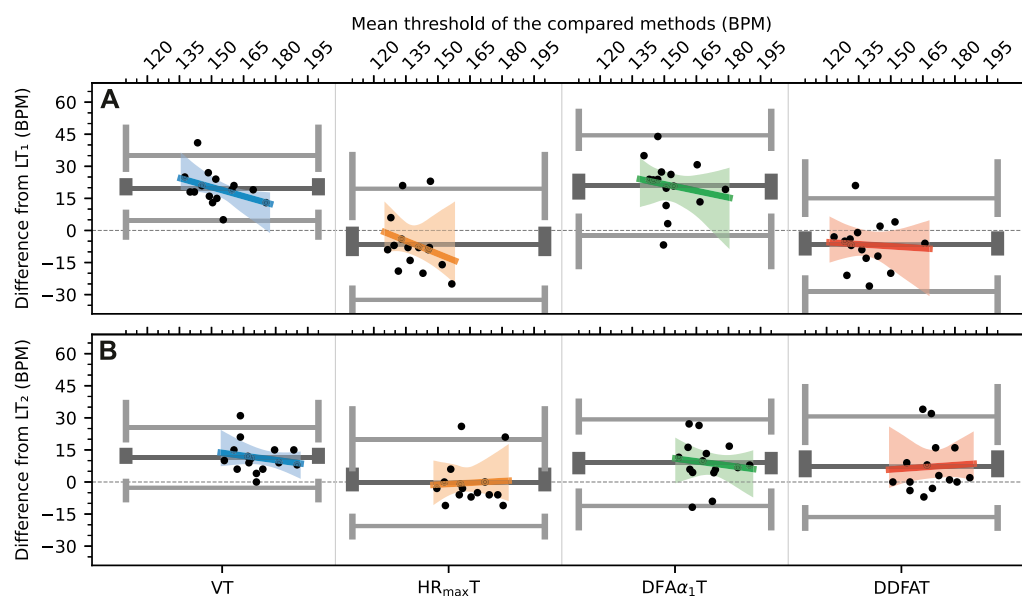
## 4 Discussion

This study explores the possibilities to determine the physiological exercise thresholds through dynamical correlation properties of RRIs captured by DDFA. We implemented a method to quantify the exercise intensity by performing the DDFA of the HR time series and calculating the thresholds from the scaling exponents  $\alpha(\text{HR}, s)$  for incremental cyclo-ergometer exercises. In accordance with previous studies (Molkkari et al., 2020), it is found that physical exercise leads to a decrease in the scaling exponents down to the regime of anticorrelations between the RRIs. These anticorrelations may reach relatively long scales comprising dozens of RRIs. A combination of the scaling exponents for scales 5–64 RRIs yields a model to predict the exercise thresholds. The overall agreement with both LTs is good, and the method does not yield notable linear bias with respect to the HR. However, more data including different testing protocols are needed for further validation. If needed, the proposed model can be modified in a straightforward way to account for, e.g., different testing protocols.

Our HR-based method is directly applicable in all wearable technologies measuring the HR. As demonstrated in Figure 1, the thresholds are computed from the RRIs dynamically, which opens the possibility of calculating the thresholds in real time with a constant data flow from the measurement device, yielding a comprehensive image of the whole exercise. Hence, we find considerable potential in the implementation of our method to the wearable devices that are intensively becoming more popular in the consumer market. Such commercial development is already foreseen (Molkkari and Räsänen, 2023).

Of other HR-based methods, HR<sub>max</sub>T compares the present HR to the expected or measured maximum HR and estimates the thresholds according to pre-determined fractions (percentages) HR/HR<sub>max</sub> for both thresholds, respectively. Thus, the validity of the method strongly depends on the assessment of HR<sub>max</sub> and especially on the validity of the aforementioned percentage for the individual. The method can be accurate for a uniform set of studied subjects with similar training background. This is reflected in our study,





**FIGURE 4**

Bland–Altman plots of the differences to lactate for each method for (A) threshold 1 (B) threshold 2. The solid lines correspond to the mean (dark) and 95% limits of agreement (light) of the distributions.

where  $HR_{max}T$  gives relatively good results for both thresholds, but—on the other hand—the cohort consists of subjects with high fitness levels compared to the general population. In practice, the parameters may need to be adjusted for different study groups (Tanaka et al., 2001). We have implemented automatization for the definition of the regression region to  $DFA\alpha_1T$  method. We also want to point out that the consideration of individual baselines in the method could improve the results, which should be considered in the future studies with  $DFA\alpha_1T$ .

## 4.1 Limitations

There are some considerable limitations in the present study. Firstly, even though the results are fairly consistent, the study population is relatively small ( $N = 15$ ), decreasing the statistical power of the analysis. Hence, as mentioned above, further studies with larger cohorts with different testing protocols are needed. In particular, it is worth studying how accurately our method works for other endurance sports such as running, swimming or rowing, or for longer exercises, where the HR is elevated and then decreased again due to breaks (e.g., due to lactate measurements) or lowered training intensity in interval training. Furthermore, as the present study focused on healthy individuals with high fitness levels, it is important to study how the physiological thresholds hold in more diverse populations with different ages, conditions and medications such as beta-blockers (BBs), which are known to alter the HRV (Niemelä et al., 1994; Pukkila et al., 2023).

Secondly, the determination of the LTs can be done in several different ways, and there is no universal golden standard of the protocol (Jamnick et al., 2020). Here, the lactate thresholds

were determined by automated algorithm according to a Finnish standard procedure with visual inspection to ascertain sensible results. In practice, the values are usually manually corrected by physician to confirm the training zones, which is prone to subjective interpretation. In this regard, it can be debated whether VTs—often determined by an average of multiple methods—provide a more reliable benchmark than LTs. Here we resorted to LTs as the benchmark, but included also VTs in the comparison.

Finally, even though the present study contains high-quality RR intervals measured using Polar H10 sensors with up to 1,000 Hz sampling rate (Schaffarczyk et al., 2022), it remains to be studied whether wearable devices without an external HR sensor can provide accurate data for our analysis. Another limitation in this study is the lack of the ECG recording, which complicates the evaluation of the nature of the outliers. Since the ECG recording is not available, we have resorted to visual examination of the samples and we removed the beats which were statistically outliers in the current context. Previously, it has been shown that sampling rates as low as 100 Hz are applicable for some HRV measures (Kwon et al., 2018). In controlled laboratory testing with standardized test procedure, some of the popular smartwatches employing photoplethysmography (PPG) have been shown to measure HRV with sufficient accuracy in rest (Shumate et al., 2021; Nuutila et al., 2022). However, there are several sources of noise in PPG measurements including, e.g., individual variations in the body temperature and composition, gender and skin tone, as well as external factors such as light and applied pressure of the device (Fine et al., 2021). Hence, in order to estimate the physiological thresholds with PPG, significant attention has to be paid to the signal quality and preprocessing. To conclude, further studies to evaluate the DDEFA-based thresholds with PPG measurements are required.

## 4.2 Conclusion

Knowledge of aerobic and anaerobic thresholds are of great importance not only for professional athletes but also for wellness-oriented consumers to optimize their training. There is a need to develop simple yet accurate estimates for the thresholds without a need for tedious respiratory gas exchange and blood lactate measurements.

Here we have introduced a computational method that estimates the both lactate thresholds (LTs) by utilizing the RR intervals during a HR measurement in an exercise setting. The training intensity corresponds to the time- and scale-dependent changes in the scaling exponent of the dynamical detrended fluctuation analysis (DDFA). This information was used to define individualized estimates for the thresholds (DDFAT<sub>1</sub> and DDFAT<sub>2</sub>). The performance of the method was compared against LTs and ventilation thresholds (VTs), as well as to two other HR-based estimates obtained for 15 participants in a incremental cyclo-ergometer test.

Our DDFAT method was found to yield a reasonable agreement with the LTs. The combined agreement with both LT<sub>1</sub> and LT<sub>2</sub> was the best of all the tested methods. The simple estimate based on HR<sub>max</sub> was found to yield relatively good results as well, but this is possibly due to the uniform fitness profile of the subjects, for which the optimized percentages of HR<sub>max</sub> are suitable.

The VTs were found to exceed LTs by  $\geq 10$  BPM in a systematic fashion. A similar effect was found with the DFA $\alpha_1$  method, which was originally constructed to assess VTs and performed well in that regard. Our statistical analysis showed that DDFATs have no linear bias in terms of the HR and thus appear as promising estimates for different fitness conditions.

In summary, DDFA provides a simple, cost-efficient, and accurate HR-based method to assess the athletic metabolic thresholds, and it could be easily integrated in wearable devices. However, it is still to be tested if the findings are applicable to larger populations or to other sports and alternative test protocols. After performing more extensive studies, we believe that the method can be successfully integrated into modern wearable devices such as smartwatches to be utilized in everyday use.

## Data availability statement

The datasets presented in this article are not readily available because the data is not available publicly due to privacy issues. The privacy policy of Kauppi Sports Coaching Ltd. does not allow sharing the data. Requests to access the datasets should be directed to KL, kimmo@kunnontestaus.fi.

## References

American College of Sports Medicine (2013). *ACSM's guidelines for exercise testing and prescription*. Lippincott Williams & Wilkins.

## Ethics statement

The studies involving humans were approved by Tampere University Hospital Ethics Committee. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

## Author contributions

MK: Writing—original draft, Writing—review and editing, Conceptualization, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization. TP: Conceptualization, Formal Analysis, Investigation, Validation, Writing—review and editing. JK: Conceptualization, Formal Analysis, Investigation, Methodology, Data curation, Writing—review and editing. MM: Formal Analysis, Methodology, Software, Conceptualization, Investigation, Visualization, Writing—review and editing. KL: Data curation, Resources, Investigation, Writing—review and editing. ER: Funding acquisition, Project administration, Supervision, Conceptualization, Investigation, Writing—review and editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The study was conducted in Tampere University as a part of the project “MoniCardi—Monitoring of Cardiac Health Made Simple” (2022–23) funded by Business Finland.

## Conflict of interest

KL is the founder and the CEO of company Kauppi Sports Coaching Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Beaver, W. L., Wasserman, K., and Whipp, B. J. (1986). A new method for detecting anaerobic threshold by gas exchange. *J. Appl. Physiology* 60, 2020–2027. doi:10.1152/jappl.1986.60.6.2020

- Binder, R. K., Wonisch, M., Corra, U., Cohen-Solal, A., Vanhees, L., Saner, H., et al. (2008). Methodological approach to the first and second lactate threshold in incremental cardiopulmonary exercise testing. *Eur. J. Cardiovasc. Prev. rehabilitation* 15, 726–734. doi:10.1097/HJR.0b013e318328304fed4
- Blain, G., Meste, O., Bouchard, T., and Berman, S. (2005). Assessment of ventilatory thresholds during graded and maximal exercise test using time varying analysis of respiratory sinus arrhythmia. *Br. J. Sports Med.* 39, 448–452. doi:10.1136/bjism.2004.014134
- Bland, J. M., and Altman, D. G. (1999). Measuring agreement in method comparison studies. *Stat. methods Med. Res.* 8, 135–160. doi:10.1177/096228029900800204
- Cannon, D. T., Kolkhorst, F. W., and Buono, M. J. (2009). On the determination of ventilatory threshold and respiratory compensation point via respiratory frequency. *Int. J. Sports Med.* 30, 157–162. doi:10.1055/s-0028-1104569
- Cottin, F., Leprière, P.-M., Lopes, P., Papelier, Y., Médigue, C., and Billat, V. (2006). Assessment of ventilatory thresholds from heart rate variability in well-trained subjects during cycling. *Int. J. Sports Med.* 27, 959–967. doi:10.1055/s-2006-923849
- Di Michele, R., Gatta, G., Di Leo, A., Cortesi, M., Andina, F., Tam, E., et al. (2012). Estimation of the anaerobic threshold from heart rate variability in an incremental swimming test. *J. Strength & Cond. Res.* 26, 3059–3066. doi:10.1519/JSC.0b013e318245bde1
- dos Santos, L., Barroso, J. J., Macau, E. E., and de Godoy, M. F. (2013). Application of an automatic adaptive filter for heart rate variability analysis. *Med. Eng. Phys.* 35, 1778–1785. doi:10.1016/j.medengphy.2013.07.009
- Efron, B. (1987). Better bootstrap confidence intervals. *J. Am. Stat. Assoc.* 82, 171–185. doi:10.1080/01621459.1987.10478410
- Faude, O., Kindermann, W., and Meyer, T. (2009). Lactate threshold concepts. *Sports Med.* 39, 469–490. doi:10.2165/00007256-200939060-00003
- Fine, J., Branan, K. L., Rodriguez, A. J., Boonya-Ananta, T., Ramella-Roman, J. C., McShane, M. J., et al. (2021). Sources of inaccuracy in photoplethysmography for continuous cardiovascular monitoring. *Biosensors* 11, 126. doi:10.3390/bios11040126
- Foster, C. (1983). VO<sub>2</sub> max and training indices as determinants of competitive running performance. *J. Sports Sci.* 1, 13–22. doi:10.1080/02640418308729657
- Gaskill, S. E., Ruby, B. C., Walker, A. J., Sanchez, O. A., Serfass, R. C., Leon, A. S., et al. (2001). Validity and reliability of combining three methods to determine ventilatory threshold. *Med. Sci. Sports Exerc.* 33, 1841–1848. doi:10.1097/00005768-200111000-00007
- Gilgen-Ammann, R., Schweizer, T., and Wyss, T. (2019). RR interval signal quality of a heart rate monitor and an ECG Holter at rest and during exercise. *Eur. J. Appl. Physiology* 119, 1525–1532. doi:10.1007/s00421-019-04142-5
- Goldberger, A. L., Amaral, L. A., Hausdorff, J. M., Ivanov, P. C., Peng, C.-K., and Stanley, H. E. (2002). Fractal dynamics in physiology: alterations with disease and aging. *Proc. Natl. Acad. Sci.* 99, 2466–2472. doi:10.1073/pnas.012579499
- Gronwald, T., and Hoos, O. (2020). Correlation properties of heart rate variability during endurance exercise: a systematic review. *Ann. Noninvasive Electrocardiol.* 25, e12697. doi:10.1111/anec.12697
- Gronwald, T., Hoos, O., Ludyga, S., and Hottenrott, K. (2019). Non-linear dynamics of heart rate variability during incremental cycling exercise. *Res. Sports Med.* 27, 88–98. doi:10.1080/15438627.2018.1502182
- Hautala, A. J., Mäkikallio, T. H., Seppänen, R. T., Huikuri, H. V., and Tulppo, M. P. (2003). Short-term correlation properties of RR interval dynamics at different exercise intensity levels. *Clin. physiology Funct. imaging* 23, 215–223. doi:10.1046/j.1475-097x.2003.00499.x
- Hoffman, R. L. (1999). Effects of training at the ventilatory threshold on the ventilatory threshold and performance in trained distance runners. *J. Strength & Cond. Res.* 13, 118–123. doi:10.1519/00124278-199905000-00004
- Jamnack, N. A., Botella, J., Pyne, D. B., and Bishop, D. J. (2018). Manipulating graded exercise test variables affects the validity of the lactate threshold and [Formula: see text]. *PLoS one* 13, e0199794. doi:10.1371/journal.pone.0199794
- Jamnack, N. A., Pettitt, R. W., Granata, C., Pyne, D. B., and Bishop, D. J. (2020). An examination and critique of current methods to determine exercise intensity. *Sports Med.* 50, 1729–1756. doi:10.1007/s40279-020-01322-8
- Jones, A. M. (1998). A five year physiological case study of an olympic runner. *Br. J. Sports Med.* 32, 39–43. doi:10.1136/bjism.32.1.39
- Karasik, R., Sapir, N., Ashkenazy, Y., Ivanov, P. C., Dvir, I., Lavie, P., et al. (2002). Correlation differences in heartbeat fluctuations during rest and exercise. *Phys. Rev. E* 66, 062902. doi:10.1103/physreve.66.062902
- Keskinen, K. L., Häkkinen, K., and Kallinen, M. (2018). *Fyysisen kunnon mittaaminen: käsi-ja oppikirja kuntotestajille*. Helsinki: Liikuntatieteellinen Seura.
- Kim, K. J., Rivas, E., Prejean, B., Frisco, D., Young, M., and Downs, M. (2021). Novel computerized method for automated determination of ventilatory threshold and respiratory compensation point. *Front. Physiology* 12, 782167. doi:10.3389/fphys.2021.782167
- Kwon, O., Jeong, J., Kim, H. B., Kwon, I. H., Park, S. Y., Kim, J. E., et al. (2018). Electrocardiogram sampling frequency range acceptable for heart rate variability analysis. *Healthc. Inf. Res.* 24, 198–206. doi:10.4258/hir.2018.24.3.198
- McPartland, D., Pree, A., Malpeli, R., and Telford, A. (2010). *Nelson physical education studies for WA 2A, 2B*. Nelson Physical Education Studies for WA (Cengage Learning Australia).
- Molkkari, M., Angelotti, G., Emig, T., and Räsänen, E. (2020). Dynamical heart beat correlations during running. *Sci. Rep.* 10, 13627. doi:10.1038/s41598-020-70358-7
- Molkkari, M., and Räsänen, E. (2023). Inter-beat interval of heart for estimating condition of subject. *Pat. pending*.
- Neves, L. N. S., Neto, V. H. G., Araujo, I. Z., Barbieri, R. A., Leite, R. D., and Carletti, L. (2022). Is there agreement and precision between heart rate variability, ventilatory, and lactate thresholds in healthy adults? *Int. J. Environ. Res. Public Health* 19, 14676. doi:10.3390/ijerph192214676
- Newell, J., Higgins, D., Madden, N., Cruickshank, J., Einbeck, J., McMillan, K., et al. (2007). Software for calculating blood lactate endurance markers. *J. Sports Sci.* 25, 1403–1409. doi:10.1080/02640410601128922
- Niemelä, M. J., Airaksinen, K., and Huikuri, H. V. (1994). Effect of beta-blockade on heart rate variability in patients with coronary artery disease. *J. Am. Coll. Cardiol.* 23, 1370–1377. doi:10.1016/0735-1097(94)90379-4
- Nuutila, O.-P., Korhonen, E., Laukkanen, J., and Kyröläinen, H. (2022). Validity of the wrist-worn Polar vantage V2 to measure heart rate and heart rate variability at rest. *Sensors* 22, 137. doi:10.3390/s22010137
- Peng, C.-K., Havlin, S., Stanley, H. E., and Goldberger, A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos Interdiscip. J. nonlinear Sci.* 5, 82–87. doi:10.1063/1.166141
- Pukkila, T., Molkkari, M., Kanniainen, M., Hernesniemi, J., Nikus, K., Lyytikäinen, L.-P., et al. (2023). Effects of beta blocker therapy on RR interval correlations during exercise. In prep.
- Ramos-Campo, D. J., Rubio-Arias, J. A., Ávila-Gandía, V., Marin-Pagán, C., Luque, A., and Alcaraz, P. E. (2017). Heart rate variability to assess ventilatory thresholds in professional basketball players. *J. Sport Health Sci.* 6, 468–473. doi:10.1016/j.jshs.2016.01.002
- Rogers, B., Giles, D., Draper, N., Hoos, O., and 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. *Front. physiology* 11, 596567. doi:10.3389/fphys.2020.596567
- Rogers, B., Giles, D., Draper, N., Mourot, L., and Gronwald, T. (2021b). Detection of the anaerobic threshold in endurance sports: validation of a new method using correlation properties of heart rate variability. *J. Funct. Morphol. Kinesiol.* 6, 38. doi:10.3390/jfkm6020038
- Schaffarczyk, M., Rogers, B., Reer, R., and Gronwald, T. (2022). Validity of the Polar H10 sensor for heart rate variability analysis during resting state and incremental exercise in recreational men and women. *Sensors* 22, 6536. doi:10.3390/s22176536
- Şekir, U., Özyener, F., and Gür, H. (2002). Effect of time of day on the relationship between lactate and ventilatory thresholds: a brief report. *J. Sports Sci. Med.* 1, 136–140.
- Shaffer, F., and Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Front. public health* 5, 258. doi:10.3389/fpubh.2017.00258
- Shookster, D., Lindsey, B., Cortes, N., and Martin, J. R. (2020). Accuracy of commonly used age-predicted maximal heart rate equations. *Int. J. Exerc. Sci.* 13, 1242–1250.
- Shumate, T., Link, M., Furness, J., Kemp-Smith, K., Simas, V., and Climstein, M. (2021). Validity of the Polar Vantage M watch when measuring heart rate at different exercise intensities. *PeerJ* 9, e10893. doi:10.7717/peerj.10893
- Stöggel, T. L., and Sperlich, B. (2019). Editorial: training intensity, volume and recovery distribution among elite and recreational endurance athletes. *Front. Physiology* 10, 592. doi:10.3389/fphys.2019.00592
- Tanaka, H., Monahan, K. D., and Seals, D. R. (2001). Age-predicted maximal heart rate revisited. *J. Am. Coll. Cardiol.* 37, 153–156. doi:10.1016/s0735-1097(00)01054-8
- Tarvainen, M., Ranta-aho, P., and Karjalainen, P. (2002). An advanced detrending method with application to HRV analysis. *IEEE Trans. Biomed. Eng.* 49, 172–175. doi:10.1109/10.979357
- Tjelta, L. I. (2019). Three Norwegian brothers all european 1500 m champions: what is the secret? *Int. J. Sports Sci. Coach.* 14, 694–700. doi:10.1177/1747954119872321
- van Campen, C. L. M., Rowe, P. C., and Visser, F. C. (2020). Heart rate thresholds to limit activity in myalgic encephalomyelitis/chronic fatigue syndrome patients (pacing): comparison of heart rate formulae and measurements of the heart rate at the lactic acidosis threshold during cardiopulmonary exercise testing. *Adv. Phys. Educ.* 10, 138–154. doi:10.4236/ape.2020.102013
- Yeh, M. P., Gardner, R. M., Adams, T., Yanowitz, F., and Crapo, R. (1983). “Anaerobic threshold”: problems of determination and validation. *J. Appl. Physiology* 55, 1178–1186. doi:10.1152/jappl.1983.55.4.1178
- Zignoli, A. (2023). Machine learning models for the automatic detection of exercise thresholds in cardiopulmonary exercising tests: from regression to generation to explanation. *Sensors* 23, 826. doi:10.3390/s23020826