Heart Rate Variability during Exercise: A Comparison of Artefact Correction Methods

David Giles[1] and Nick Draper[2]

[1]Department of Life Sciences, College of Life and Natural Sciences, University of Derby, UK

[2]School of Sport and Physical Education, University of Canterbury, Christchurch, NZ

Abstract: There is a need for standard practice in the collection and processing of RR interval data recorded using heart rate monitors (HRM) in research. This paper assessed the validity of (a) RR intervals and heart rate variability (HRV) data obtained using a HRM during incremental exercise, and (b) artefact correction methods. Eighteen participants completed an active orthostatic test and incremental running $\dot{V}O_{2MAX}$ test, while simultaneous recordings using the V800 and an electrocardiogram were made. Artefacts were corrected by deletion; degree zero, linear, cubic and spline interpolation; and using Kubios HRV software. Agreement was assessed using percentage bias and effect size (ES), intra-class correlation coefficients (ICC) and Bland–Altman limits of agreement (LoA). The number of artefacts increased relative to the intensity of the exercise, to a peak of 4.46% during 80-100% $\dot{V}O_{2MAX}$. Correction of RR intervals was necessary with unacceptably increased percentage bias, LoA and ES and reduced ICC in all but supine and standing recordings. All correction methods resulted in data with reduced percentage bias and ES for resting and <60% $\dot{V}O_{2MAX}$ exercise recordings. However, at >60% $\dot{V}O_{2MAX}$, even when correction methods were applied, large amounts of variation were present in RMSSD, LF:HF ratio, SD1 and SampEn. Linear interpolation produced corrected RR intervals with the lowest bias and ES. However, caution should be given to HRV parameters at high exercise intensities, as large amounts of variation were still present. Recommendations for minimising recording artefacts are discussed, along with guidelines for their identification, correction and reporting.

Keywords: Heart Rate Variability, Artefacts, Exercise, Time domain analysis, Frequency domain analysis, Non-linear analysis

INTRODUCTION

Heart rate variability (HRV) analysis quantifies fluctuations in the duration of interbeat intervals (29). Ideally, HRV analysis is performed on RR intervals that contain only pure sinus beats (normal to normal intervals; NN). However, both technical and biological artefacts occur in the data (15-17). If artefacts are left unedited, they introduce bias and present a significant problem for the interpretation of calculated time, frequency and non-linear parameters (23); as such, pre-processing is necessary. Pre-processing of RR interval data involves the editing of

artefacts by deletion, interpolation or filtering. However, editing methods have distinct effects on the RR interval time series and HRV results. If appropriate corrections are performed the presence of artefacts should have a less dramatic influence on HRV data (27).

Consideration of the validity of RR interval data recorded using consumer heart rate monitors (HRM) before analysis is essential, especially since they are increasingly being used in research (11, 31, 35). Their portability has enabled the recording of RR data in situations where it was not previously possible and not conceivable with lab-based electrocardiograms (ECG), or even ambulatory ECGs (21, 22). Finally, and most significantly, previous research has demonstrated the validity of HRMs at rest, finding good agreement with a simultaneously recorded ECG with small but acceptable levels of variation (e.g. Polar V800 (17), RS800 (25, 33), S810 (15, 16, 32, 35)). The validity of recorded RR intervals holds true as long as artefacts are corrected before analysis. While it is not possible to determine the source of artefacts in HRMs signal; it is thought that they result because of a decrease in contact between the skin and the electrode, causing a reduction in the amplitude of the detected R-wave (16, 17).

Artefacts require correction even at rest when their occurrence is minimal (15-17). Simply discarding erroneous sections of data, especially sections that involve more movement and potentially a greater number of artefacts, leads to bias (9). However, the methods that are used for the correction of artefacts in recent literature vary considerably. While several studies have used various interpolation techniques (8, 12), others simply deleted offending intervals (20) or, most frequently, do not mention correction at all (13, 14, 31). Equally, research that has specifically examined the validity of HRMs have used several methods (4, 32). As such, there is a need for standard practice in the collection and processing of RR interval data recorded using HRMs in research as, currently, there are no standard criteria.

The present study aimed to develop on existing criteria for the identification of errors and determine the validity of correction methods when compared to a simultaneously recorded ECG. Secondly, the study aimed to identify alterations in the occurrence of artefacts during incremental exercise. While there are more complex methods for the correction of artefacts, the present study focused on accessibility, ease of use, and validity.

MATERIALS AND METHODS

Experimental Approach to the Problem

RR interval data was simultaneously recorded using the Polar V800 and a three-lead ECG at rest and during incremental exercise. Artefacts were: (1) identified following the proposed guidelines of Giles, Draper and Neil

(17) developed from Gamelin, Berthoin and Bosquet (16); (2) artefacts in the HRM RR interval time series were corrected by deletion, interpolation and using Kubios HRV; (3) time, frequency and nonlinear HRV parameters were calculated for ECG RR intervals, uncorrected HRM interval and for each correction methods; (4) agreement was assessed using percentage bias and effect size (ES), intra-class correlation coefficients (ICC) and Bland–Altman limits of agreement (LoA).

Subjects

Eighteen male volunteers (age 27.6 ± 9.9 years; height 1.77 ± 0.09 metres; mass 77.0 ± 7.9 Kg) agreed to participate in the study. No participant was known to be taking medication or have any cardiovascular or respiratory diseases or illnesses that may have influenced the procedures carried out. Participants completed written informed consent and medical health questionnaires. Approval for the study was granted by the Universities Ethics Committee [LSREC_1415_16] and conformed to the principles of the Declaration of Helsinki.

Procedures – RR interval recording

Participants were asked to abstain from caffeine-containing food and drink for 24-hours before the test. Participant's skin was cleaned (shaved if necessary) and prepared for the attachment of the ECG electrodes. The electrodes were placed in a CM5 configuration (right fifth interspace, manubrium and left fifth interspace, ensuring that they did not interfere with the fit of the HRM strap (Polar H7). The electrode belt was dampened and placed following Polar's guidelines, tightly but comfortably just below the chest muscles.

Resting measurements. Resting measurements were conducted in two positions, supine and, following an active orthostatic challenge, standing. Participants were kept in silence, without moving, and with limited audible and visual stimulation. Recordings lasted for 10 minutes in the supine position and 7 minutes in standing position. No attempt was made to control participant's respiratory frequency or tidal volume (24). Environmental conditions were maintained at 20.6 ± 1.0 °C and 50% relative humidity.

Exercise Measurement. Participants performed an incremental $\dot{V}O_{2MAX}$ test, following a modified Bruce Protocol, on a motor-driven treadmill (Woodway; Birmingham, UK). The treadmill was programmed for increases in angle of inclination and speed from 2.7 km/h at 10% grade, increasing by 1.3 km/h (0.6 mph) and 2% grade every three minutes until volitional exhaustion. Oxygen kinetics were recorded continuously using a breath-to-breath ergo spirometry system (Metalyzer 3B, Cortex Biophysik GmbH).

Data Recording. RR interval data was recorded simultaneously using a V800 Polar HRM with a Polar H7 chest strap and a three lead ECG (MP36, Biopac Systems Ltd.), at a sampling frequency of 1000 Hz for both devices. R-wave peaks from the ECG were detected automatically using a custom peak detection algorithm in

Matlab (Mathworks, Cambridge). The raw ECG traces and detected R-waves were manually assessed to ensure that they had been correctly detected. Ectopic beats were noted, but not corrected at that stage of analysis.

RR Interval Data Handling. Data was saved as RR interval data files, with intervals in milliseconds (ms). For the Polar HRM raw, unfiltered RR data was exported from the Polar Flow web service as a space delimited .txt file. Both the ECG and HRM raw RR signal start points were manually matched before further analysis. Interval data was divided into the following sections: Supine, Standing, $< 40\%$ $\dot{V}O_{2MAX}$, $40 - 60\%$ $\dot{V}O_{2MAX}$, $60 - 80\%$ $\dot{V}O_{2MAX}$, $80 - 100\%$ $\dot{V}O_{2MAX}$, and active recovery (walking).

Procedures – Artefact handling and correction procedure

Identification of artefacts. The ECG and Polar V800 interval time series were compared side-by-side to identify artefacts and ectopic beats greater than 20 ms (Fig 1). Ectopic beats are not easily differentiated from artefacts in HRM RR signals; ectopic beats present identically to the artefacts listed in Table I (usually as a Type 4, or Type 1), and should be corrected as such. The methods proposed for the correction of artefacts and ectopic beats are identical to those used in previously in replacement of ectopic beats in ECG signals (23, 26). For the purpose of the present paper, non-sinus beats (0.0015% RR intervals) were replaced during analysis in both signals with interpolated data from adjacent RR intervals. While this does not represent real-world use, it ensured that bias from the ectopic beats was not introduced. Before correction discrepancies between the two signals were identified and synchronicity maintained with the insertion of a 0 ms interval.

Table I: Classification of measurement artefacts (15-17)

Type of Artefact	Correctable?	Description
T1	N	A discrepancy greater than 20 ms at a single interval, either positive or negative.
T2	Y	A long interval, followed by a short interval. While the two points either side were unaffected.
T3	Y	Short interval, followed by a long interval. While the two points either side were unaffected.
T4	Y	Missed interval(s) on the HRM, equivalent to two or three ECG RR intervals.
T5	Y	Extra, short, RR intervals from the HRM, in the space of one on the ECG.
T6-a	N	RR interval(s) entirely missed by the HRM, undetectable.
T6-b	Y	RR interval(s) entirely missed by the HRM, detectable.

Notes: T Error Type; ms millisecond; HRM heart rate monitor; ECG electrocardiogram.
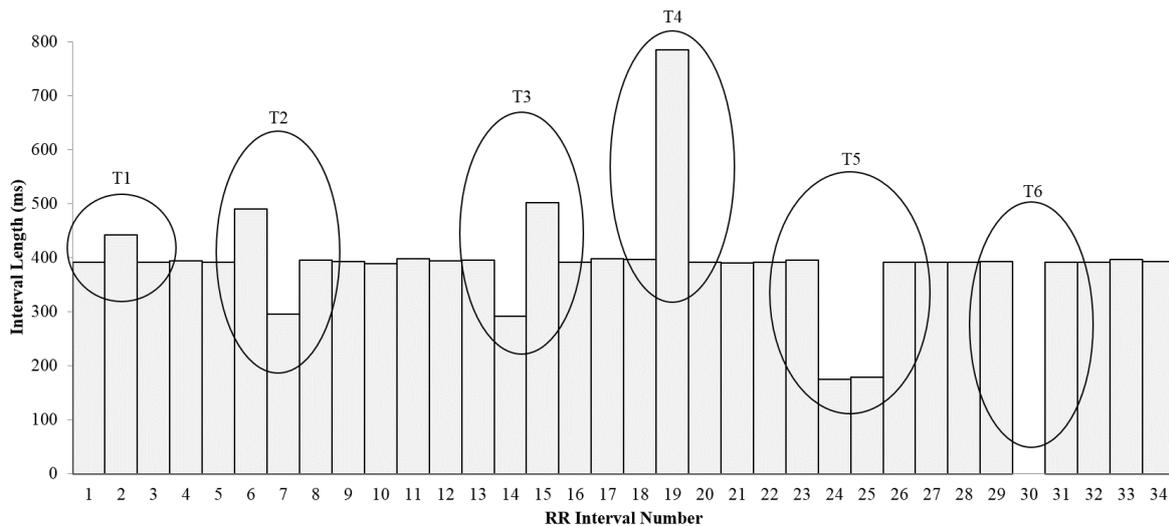
Fig 1: Example RR interval data, with examples of each of the 7 types of artefacts highlighted

Categorisation of artefacts. Following visual identification, discrepancies were assigned to one of the six types of artefacts given in Table I. The artefacts identified are based on the research of Giles, Draper and Neil (17) developed from Gamelin, Berthoin and Bosquet (16). The frequency of the occurrence of each of the artefacts was recorded.

Methods for correction. Once identified, artefacts in the RR time series were corrected. Previously, Gamelin, Baquet, Berthoin and Bosquet (15) corrected all artefacts detected (T1-T5), however, this does not represent typical use of the HRM, as, without the utilisation of a simultaneous ECG recording it is not possible to detect T1 artefacts that are less than a 50% increase in length over the previous interval. Equally, T6-a is undetectable without a simultaneous recording. Giles, Draper and Neil (17) proposed the correction of unidentifiable artefacts (short T1 and T6-b), whilst still correcting those identifiable without a simultaneous ECG recording (T1-T5 and T6-b) following the guidelines given in Table I. The exception for the application of correction methods was T5, which were simply summed in all correction methods. The following correction methods were selected, based on their applicability to RR interval signals recorded by HRMs and the ease of their application to the RR interval data. When selecting methods for the identification and correction of artefacts consideration was given to 1) the ease of applying the identification and correction methods to data, ensuring that they do not require preparatory software or specific programming knowledge; 2) do not need a simultaneous ECG for their identification; 3) are applicable to RR interval recordings completed at rest and during exercise 4) and, finally, are easy to interpret.

The methods used for artefact correction were: **(a)** *Uncorrected*, no correction applied to any intervals; data was left as recorded. **(b)** *Deletion*, erroneous RR interval(s) were simply deleted from the time series. Deletion can have a significant effect on HRV parameters due to changes in the length of the signal, particularly in short-

term recordings and frequency domain parameters (23, 26). Deletion introduces step-like shapes into the RR interval time series resulting in changes in variability as well as decreasing the length of the signal, producing false LF and HF frequency components (26). **Interpolation.** Interpolation methods, in contrast to the deletion, replaces erroneous non-normal RR intervals with interpolated intervals. Critically, interpolation allows for the length of the recording to remain the same, mitigating the issue of reduced signal length. **(c)** *Degree Zero (average)*, substitution of artefact(s) with a mean value that is calculated from surrounding RR intervals. On longer sections of artefacts, degree zero interpolation results in the same averaged value over a whole segment, resulting in a flat shape, introducing false trends and increasing LF and VLF power (5, 26). **(d)** *Degree One (linear)*, a straight line is fitted over the irregular intervals to obtain new values. As with degree zero, on longer sections of artefacts slope like shapes occur, introducing false trends and potentially increasing LF and VLF components (5, 26). **(e)** *Cubic*, cubic interpolation uses four data points to compute the polynomial; there are no constraints on the derivatives. Cubic interpolation does not result in flat sections of data. However, as non-linear analysis is concerned with the complexity and irregularity of heartbeat series the introduction of a potentially falsely correlated signal is of concern, particular if there is a significant number of erroneous intervals (23, 26). **(f)** *Spline*, cubic spline interpolation. Smooth values are estimated through a number of data points by fitting a third-degree polynomial. Cubic spline interpolation computes a third order polynomial from only two data points with the additional constraint that the first and second derivative at the interpolation points are continuous. As with cubic interpolation, spline interpolation may also introduce false correlations into the signal. **(g)** *Kubios* HRV software, Kubios software provides options for the detection and correction of ectopic beats and artefacts. Ectopic beat free RR interval data was entered into Kubios; the appropriate correction level was chosen by visual inspection of the interval tacogram (the threshold for detecting artefact beats), to ensure that the correction removed the artefacts but did not distort normal RR intervals. When the correction was applied, detected artefact beats were automatically replaced using cubic spline interpolation (28).

Procedures – Calculation of HRV parameters and statistical analysis

Following the analysis of the Polar RR trace for artefacts and the replacement of ectopic, erroneous and noisy complexes, the RR interval data was considered normal, and thus described as NN data. All HRV analysis was performed with the freely available Kubios HRV software (Version 2.2; 28).

Heart rate variability parameters. For the calculation of HRV all intervals from each section were used for ECG and corrected HRM recordings. These selected segments were analysed for time, frequency domain and nonlinear components. Not all parameters are relevant to exercise, or the length of the recording but are presented

for clarity and completeness. (a) Time domain analysis concerns the statistical representation of the variation in NN intervals within the sample (19). A number of parameters were calculated: SDNN is the standard deviation of the NN intervals, RMSSD the root mean squared of the successive difference of intervals and pNN50% the number of successive intervals that differ by more than 50 ms, expressed as a % of the total (19). (b) Frequency domain analysis allows for the identification of sympathetic and parasympathetic contributions of HRV. The non-parametric power spectral density (PSD) analysis provides basic information on how power and, therefore, the variance, distributes as a function of frequency using a fast Fourier transformation. A fast Fourier transformation allows the analysis of the components of the power spectrum density to be quantified into different frequency bands (1). Three spectral components were calculated, Very Low Frequency (VLF; 0.00 - 0.04 Hz), Low Frequency (LF; 0.04 - 0.15 Hz) and High Frequency (HF; 0.15 - 0.40 Hz). Additionally, normalised LF and HF power were calculated (as a percentage of the sum of LF and HF power) and the ratio LF:HF power. (c) Non-linear analysis, given the complex control systems of the heart it is reasonable to assume nonlinear mechanisms are involved in the genesis of HRV; non-linear analysis of NN intervals describes the chaotic nature of the signal (28). The values SD1 and SD2 were calculated from a Poincaré Plot of the data, by fitting an ellipse oriented according to the line-of-identity and computing the standard deviation of the points perpendicular (SD1) to and along the line-of-identity (SD2) (7). The analysis comprised of Sample Entropy was also calculated, measuring the complexity of the NN series, low entropy arises from extremely regular time series, higher values reflect more complexity, and the highest values are typical for stochastic datasets (34).

Statistical analysis

Descriptive statistics were first calculated for all variables; all values are reported as Mean ± SD. Normal distribution and homogeneity of variance were assessed by visual inspection of the frequency histogram and with either a Kolmogorov-Smirnov (RR intervals) or Shapiro-Wilk test (HRV parameters) depending on the number of samples. Homoscedasticity was determined through the analysis of the plot of the standardised residuals. A one-way repeated measures ANOVA was used to assess changes in the frequency of artefacts. Percentage bias was calculated between each of the correction methods and the ECG RR intervals (mean difference ÷ ECG component mean x 100). The magnitude of the difference of the RR intervals and the HRV parameters was calculated by determining the effect size (ES) which represents the mean difference over the standard deviation of the difference (30); the difference was considered small when ES ≤ 0.2, moderate when ES ≤ 0.5, and great when ES > 0.8 (10). Relative reliability was assessed for all variables by calculating the intraclass correlation

coefficient (ICC) (36), model 3.1 was used (3). Bland–Altman 95% limits of agreement (LoA) were calculated for all RR and HRV parameters (6). If heteroscedasticity was present in any HRV data, it was log-transformed before the calculation of the LoA. The level for accepting the statistical significance of tests was set at $p < 0.05$ for all analysis. All data was analysed using SPSS (Version 22; Chicago, IL).

RESULTS

Alterations in select time, frequency and nonlinear HRV parameters are provided in Table II for supine and standing resting measurements, exercise and active recovery.

Table II: Alterations in select time, frequency and nonlinear HR variability parameters at rest, during incremental exercise and active recovery

	Supine	Standing	<40%	40 - 60%	60 - 80%	80 - 100%	Active Recovery
Mean RR (ms)	1046.7 ± 168.2	762.7 ± 117	638.3 ± 95.7	513.9 ± 85.5	406.4 ± 48.7	352.3 ± 43.9	401.7 ± 56.8
SDNN (ms)	81.3 ± 40.6	65.1 ± 18.2	58.3 ± 28.2	40.3 ± 16.6	27.4 ± 13.1	20.5 ± 19.9	54.0 ± 36
RMSSD (ms)	75.9 ± 50.1	31.9 ± 13.9	22.2 ± 12.5	8.9 ± 9.0	3.2 ± 1.1	3.4 ± 1.1	5.6 ± 7.2
pNN50 (%)	38.9 ± 26.5	8.6 ± 7.9	4.7 ± 6.2	0.5 ± 1.7	0.0 ± 0.0	0.0 ± 0.0	0.5 ± 1.9
VLF Power (ms^2)	2045.6 ± 1795.7	1176.3 ± 747.4	1484.4 ± 2438.1	773.6 ± 838.2	279.7 ± 222.9	199.0 ± 433.0	1527.0 ± 3267.0
LF Power (ms^2)	1423.7 ± 1291.4	1707.4 ± 1290.4	863.1 ± 1067.9	245 ± 763.2	3.7 ± 3.1	3.0 ± 9.3	183.0 ± 674.9
HF Power (ms^2)	3978.2 ± 4800.4	599.3 ± 489.7	332.6 ± 401.4	92.4 ± 284.2	1.8 ± 2.6	0.9 ± 2.0	53.3 ± 200.5
nuLF Power	36.8 ± 16.4	74.9 ± 9.1	71.5 ± 13.7	79 ± 12.7	70.9 ± 23.5	54.0 ± 20.5	80.5 ± 11.4
nuHF Power	63.0 ± 16.4	25.0 ± 9.1	28.4 ± 13.7	20.9 ± 12.7	28.5 ± 23	45.6 ± 20.4	19.4 ± 11.3
LF:HF Ratio	0.8 ± 0.9	3.8 ± 2.9	3.4 ± 2.4	6.9 ± 7.6	5.6 ± 5.5	1.8 ± 1.5	5.9 ± 3.9
SD1	53.7 ± 35.5	22.6 ± 9.8	15.7 ± 8.9	6.3 ± 6.4	2.3 ± 0.8	2.4 ± 0.8	4.0 ± 5.1
SD2	100.9 ± 46.9	88.9 ± 24.3	80.0 ± 39.2	56.4 ± 23	38.5 ± 18.5	28.8 ± 28.1	76.1 ± 50.6
Sample Entropy	1.5 ± 0.3	0.9 ± 0.3	0.8 ± 0.4	0.4 ± 0.3	0.3 ± 0.2	0.6 ± 0.4	0.2 ± 0.2

Note: SDNN Standard Deviation of NN intervals; RMSSD Root Mean Squared of the Successive Difference of Intervals; pNN50% Percentage of successive intervals that differ by more than 50 ms; VLF Very Low Frequency; LF Low Frequency; HF High Frequency; nuLF Normalised Low Frequency Power; nuHF Normalised High Frequency Power; LF: HF Ratio; SD 1 and 2 standard deviation of the points

The number of RR intervals detected (59545 total) are shown in Table III, along with the type and frequency of artefacts identified. Type 4 made up 96.8% of detected artefacts, Type 6a 1.4%, Type 3 1% and the remaining artefacts <1%. No Type 2 artefacts were detected. The frequency of the occurrence of artefacts increased significantly with the rise in the exercise intensity (ANOVA, $p > 0.005$).

Table III: Types of error at each of the exercise intensities (Supine, Standing, <40%, 40-60%, 60-80%, 80-100% $\dot{V}O_{2MAX}$ and active recovery)

				Number of Artefacts					
Type of Error	**Supine**	**Standing**	**<40%**	**40-60%**	**60-80%**	**80-100%**	**Active Recovery**	**TOTAL**	**%**
1	2	1	0	0	2	1	0	6	0.5%
2	0	0	0	0	0	0	0	0	0.0%
3	0	0	2	1	3	5	1	12	1.0%
4	6	1	49	125	241	519	205	1146	96.8%
5	0	0	3	0	0	0	0	3	0.3%
6a	2	2	2	3	4	2	1	16	1.4%
6b	0	1	0	0	0	0	0	1	0.1%
ECTOPIC	1	1	3	2	19	41	18	85	
TOTAL INTERVALS	10586	9579	6547	6829	7596	12057	6351	59545	

Note: $\dot{V}O_{2MAX}$ maximum aerobic capacity

Along with the increased number of artefacts, uncorrected HRM RR intervals saw increased percentage bias, LoA and ES with an increase in exercise intensity. Except supine and standing recordings (0.09% and 0.05%, respectively), the number of artefacts resulted in moderate to large effect size and increased bias (Table IV), demonstrating the need for correction in all recordings, except those with less than ~ 0.1% artefacts. All correction methods resulted in small percentage bias and ES for intervals recorded during resting and <60% $\dot{V}O_{2MAX}$ (Table IV). Linear interpolation produced corrected intervals with the lowest bias and ES. However, values of RMSSD, LF:HF ratio, SD1 and SampEn at 60+ $\dot{V}O_{2MAX}$ exercise intensities showed large bias and ES, and increased LoA and reduced ICC, regardless of correction (See Tables in Supplemental Digital Content for raw ICC and LoA for each of the correction methods at each exercise intensity).

Table IV: Comparison of HR variability parameters from differing RR interval correction methods between the ECG and Polar V800 HRM, with percentage bias (%) and effect size (ES)

		PERCENTAGE BIAS							EFFECT SIZE						
		UC	DEL	MEAN	LIN	CUB	SPL	KUBIOS	UC	DEL	MEAN	LIN	CUB	SPL	KUBIOS
SDNN (ms)	SUP	3.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.07	0.00	0.00	0.00	0.00	0.00	0.00
	STA	2.8%	-0.1%	-0.1%	-0.1%	-0.1%	-0.1%	-0.1%	0.09	0.00	0.00	0.00	0.00	0.00	0.00
	<40%	**34.6%**	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	**0.66**	0.00	0.00	0.00	0.00	0.00	0.00
	40-60%	**81.1%**	0.5%	0.0%	0.1%	0.1%	0.1%	-3.2%	**1.05**	0.01	0.00	0.00	0.00	0.00	0.08
	60-80%	**152.2%**	0.0%	0.2%	0.2%	0.2%	0.2%	7.7%	**1.51**	0.00	0.00	0.00	0.00	0.00	0.15
	80-100%	**293.7%**	0.3%	0.4%	0.3%	0.4%	-0.7%	1.1%	**1.61**	0.00	0.00	0.00	0.00	0.01	0.01
	ACT REC	**83.3%**	0.1%	0.1%	0.0%	0.0%	0.0%	0.1%	**1.10**	0.00	0.00	0.00	0.00	0.00	0.00
RMSSD (ms)	SUP	7.8%	0.1%	0.0%	0.0%	0.0%	0.0%	0.1%	0.12	0.00	0.00	0.00	0.00	0.00	0.00
	STA	**15.3%**	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	**0.23**	0.00	0.00	0.00	0.00	0.00	0.00
	<40%	**188.9%**	1.1%	0.1%	0.5%	0.4%	0.4%	-2.6%	**1.17**	0.02	0.00	0.01	0.01	0.01	0.04
	40-60%	**747.0%**	4.0%	1.9%	2.6%	2.4%	2.4%	1.2%	**1.43**	0.04	0.02	0.03	0.02	0.02	0.01
	60-80%	**2468.0%**	**17.6%**	**12.5%**	**14.4%**	**13.9%**	**14.1%**	**22.1%**	**2.09**	**0.53**	**0.38**	**0.44**	**0.42**	**0.43**	**0.65**
	80-100%	**2857.8%**	**17.1%**	**10.1%**	**12.8%**	**14.9%**	**14.6%**	**20.8%**	**2.13**	**0.52**	**0.30**	**0.38**	**0.43**	**0.45**	**0.60**
	ACT REC	**1757.7%**	8.6%	4.7%	6.1%	5.7%	6.0%	7.7%	**2.28**	0.07	0.04	0.05	0.05	0.05	0.06
VLF Power (ms²)	SUP	3.8%	0.2%	0.0%	0.0%	0.0%	0.0%	-0.1%	0.04	0.00	0.00	0.00	0.00	0.00	0.00
	STA	26.5%	-0.7%	-0.6%	-0.5%	-0.5%	-0.5%	-0.6%	**0.25**	0.01	0.01	0.01	0.01	0.00	0.01
	<40%	25.0%	0.4%	0.1%	0.4%	0.4%	0.4%	-1.3%	**0.15**	0.00	0.00	0.00	0.00	0.00	0.01
	40-60%	95.1%	1.5%	-0.2%	0.0%	0.0%	0.0%	-5.3%	**0.54**	0.01	0.00	0.00	0.00	0.00	0.05
	60-80%	214.7%	-1.0%	0.0%	0.0%	0.0%	0.0%	19.5%	**0.96**	0.01	0.00	0.00	0.00	0.00	**0.20**
	80-100%	660.6%	0.6%	-0.1%	0.0%	0.0%	-0.5%	-0.5%	**0.75**	0.00	0.00	0.00	0.00	0.00	0.00
	ACT REC	89.1%	1.5%	0.1%	0.0%	0.0%	0.0%	0.0%	**0.41**	0.01	0.00	0.00	0.00	0.00	0.00
LF Power (ms²)	SUP	10.7%	0.3%	0.1%	0.1%	0.1%	0.1%	0.1%	0.11	0.00	0.00	0.00	0.00	0.00	0.00
	STA	38.1%	0.4%	0.2%	0.3%	0.3%	0.3%	0.2%	**0.27**	0.01	0.00	0.00	0.00	0.00	0.00
	<40%	144.4%	0.2%	0.1%	0.0%	0.0%	0.0%	-4.2%	**0.78**	0.00	0.00	0.00	0.00	0.00	0.03
	40-60%	1026.4%	1.4%	0.2%	0.4%	0.4%	0.4%	-0.2%	**0.65**	0.00	0.00	0.00	0.00	0.00	0.00
	60-80%	**33560.9%**	11.8%	4.5%	0.5%	0.6%	0.7%	29.3%	**1.26**	0.13	0.05	0.01	0.01	0.01	**0.31**
	80-100%	**96043.9%**	4.1%	3.2%	-1.3%	-1.1%	-0.6%	18.2%	**0.78**	0.01	0.01	0.00	0.00	0.00	0.06
	ACT REC	**1580.1%**	1.2%	0.0%	0.0%	0.0%	0.0%	0.2%	**1.13**	0.00	0.00	0.00	0.00	0.00	0.00
HF Power (ms²)	SUP	4.2%	0.0%	0.1%	0.1%	0.1%	0.1%	0.1%	0.04	0.00	0.00	0.00	0.00	0.00	0.00
	STA	75.3%	-0.7%	-0.7%	-0.6%	-0.6%	-0.6%	-0.7%	**0.32**	0.01	0.01	0.01	0.01	0.01	0.01
	<40%	**615.4%**	-0.2%	-0.2%	0.3%	0.3%	0.4%	-6.8%	**0.85**	0.00	0.00	0.00	0.00	0.00	0.06
	40-60%	**3441.4%**	1.1%	-0.3%	-0.4%	-0.4%	-0.3%	-0.2%	**0.94**	0.00	0.00	0.00	0.00	0.00	0.00
	60-80%	**151876.1%**	23.0%	20.1%	14.6%	15.7%	16.3%	36.1%	**1.22**	0.15	0.13	0.10	0.10	0.11	**0.23**
	80-100%	**490566.8%**	23.1%	22.8%	9.3%	10.3%	11.8%	141.0%	**0.81**	0.10	0.10	0.04	0.04	0.05	**0.42**
	ACT REC	**8355.0%**	0.7%	0.2%	-0.1%	-0.1%	-0.1%	0.8%	**1.51**	0.00	0.00	0.00	0.00	0.00	0.00
LF:HF Ratio	SUP	4.9%	3.3%	3.2%	3.1%	3.1%	3.1%	3.2%	0.04	0.03	0.03	0.03	0.03	0.03	0.03
	STA	-2.8%	1.1%	1.0%	1.0%	1.0%	1.0%	1.0%	0.04	0.01	0.01	0.01	0.01	0.01	0.01
	<40%	**-48.9%**	0.2%	3.1%	-0.4%	-0.6%	-0.7%	7.4%	**0.87**	0.00	0.04	0.01	0.01	0.01	0.10
	40-60%	**-39.2%**	**-10.2%**	-9.4%	-7.8%	-8.0%	-8.0%	-13.0%	0.35	0.09	0.09	0.07	0.07	0.07	0.12
	60-80%	**-73.9%**	**-34.2%**	**-36.3%**	**-22.7%**	**-24.0%**	**-24.8%**	**-36.8%**	**0.97**	**0.44**	**0.47**	**0.28**	**0.29**	**0.31**	**0.48**
	80-100%	**-58.3%**	**-19.4%**	**-26.3%**	**-14.7%**	**-16.1%**	**-16.1%**	**-37.4%**	**0.89**	**0.26**	**0.35**	**0.19**	**0.21**	**0.21**	**0.53**
	ACT REC	**-71.7%**	**-21.4%**	**-23.0%**	**-14.3%**	**-14.4%**	**-14.4%**	**-32.1%**	**1.20**	**0.35**	**0.39**	**0.23**	**0.23**	**0.24**	**0.54**
SD1	SUP	7.8%	0.1%	0.0%	0.0%	0.0%	0.0%	0.1%	0.12	0.00	0.00	0.00	0.00	0.00	0.00
	STA	**15.3%**	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	**0.23**	0.00	0.00	0.00	0.00	0.00	0.00
	<40%	**188.8%**	1.1%	0.1%	0.5%	0.4%	0.4%	-2.6%	**1.17**	0.02	0.00	0.01	0.01	0.01	0.04
	40-60%	**746.1%**	4.0%	1.9%	2.6%	2.4%	2.4%	1.2%	**1.43**	0.04	0.02	0.03	0.02	0.02	0.01
	60-80%	**2460.0%**	**17.6%**	**12.4%**	**14.3%**	**13.9%**	**14.0%**	**22.0%**	**2.09**	**0.53**	**0.38**	**0.43**	**0.42**	**0.42**	**0.64**
	80-100%	**2854.7%**	**17.1%**	**10.1%**	**12.8%**	**14.9%**	**14.6%**	**20.7%**	**2.13**	**0.52**	**0.30**	**0.38**	**0.43**	**0.44**	**0.59**
	ACT REC	**1751.0%**	8.6%	4.7%	6.0%	5.7%	5.9%	7.7%	**2.28**	0.07	0.04	0.05	0.04	0.05	0.06
SD2	SUP	2.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.05	0.00	0.00	0.00	0.00	0.00	0.00
	STA	1.6%	-0.1%	-0.1%	-0.1%	-0.1%	-0.1%	-0.1%	0.06	0.00	0.01	0.00	0.00	0.00	0.00
	<40%	**21.9%**	0.2%	0.0%	0.0%	0.0%	0.0%	0.1%	**0.44**	0.00	0.00	0.00	0.00	0.00	0.00
	40-60%	**51.7%**	0.5%	0.0%	0.0%	0.0%	0.0%	-3.2%	**0.88**	0.01	0.00	0.00	0.00	0.00	0.08
	60-80%	**99.6%**	-0.1%	0.1%	0.1%	0.1%	0.1%	7.6%	**1.27**	0.00	0.00	0.00	0.00	0.00	0.15
	80-100%	**208.5%**	0.0%	0.2%	0.1%	0.1%	-0.9%	0.9%	**1.40**	0.00	0.00	0.00	0.00	0.01	0.01
	ACT REC	**51.1%**	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	**0.73**	0.00	0.00	0.00	0.00	0.00	0.00
Sample Entropy	SUP	-1.6%	-0.1%	-0.1%	0.0%	-0.1%	-0.1%	-0.1%	0.08	0.01	0.01	0.00	0.00	0.00	0.01
	STA	-0.6%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.02	0.01	0.01	0.01	0.01	0.01	0.01
	<40%	**-23.3%**	0.3%	-1.1%	0.1%	-0.1%	-0.3%	-5.2%	**0.53**	0.01	0.02	0.00	0.00	0.01	0.11
	40-60%	**-23.7%**	9.8%	6.2%	8.0%	7.9%	7.9%	**10.9%**	**0.46**	0.18	0.11	0.14	0.14	0.14	0.20
	60-80%	**-29.1%**	**31.8%**	**21.3%**	**26.9%**	**26.8%**	**27.6%**	**21.5%**	0.37	**0.38**	0.27	**0.33**	**0.33**	**0.34**	0.26
	80-100%	**-59.2%**	**20.4%**	8.9%	**15.1%**	**15.7%**	**20.4%**	**17.2%**	**1.21**	**0.30**	0.13	**0.22**	**0.23**	**0.30**	0.26
	ACT REC	3.9%	**21.3%**	**16.1%**	**18.5%**	**19.0%**	**20.1%**	**21.8%**	0.04	**0.22**	0.17	0.19	0.20	**0.21**	**0.22**

DISCUSSION

The present study aimed to establish differences in the occurrence of artefacts in RR intervals during exercise and to determine the validity of methods for their correction. Briefly, artefacts increased relative to exercise intensity, to a peak of 4.46% during recordings made at 80-100% $\dot{V}O_{2MAX}$. Artefact correction was necessary, with large percentage bias and ES of HRV parameters in all but supine and standing recordings; correction resulted in reduced bias and ES for resting and <60% $\dot{V}O_{2MAX}$ recordings with all methods. Linear interpolation consistently produced HRM RR intervals with the lowest percentage bias and ES. However, caution should be given to the interpretation of RMSSD, LF:HF ratio, SD1 and SampEn at high (60%+ $\dot{V}O_{2MAX}$) exercise intensities, as, even when correction methods were applied, large amounts of bias was still present.

It is not possible to determine the source of the detected artefacts they may result because of increased activity of chest muscles (respiration and arm movement) and repetitive vertical oscillations causing interference and loss, or decrease, in contact between the skin and the chest strap electrode, reducing the amplitude of the detected R-wave (16, 17). Exercise that does not introduce either of these factors are unlikely to influence dramatically the number of artefacts detected. As the majority of recent HRM validation studies were conducted either at rest or during an exercise that did not involve upper body movement (e.g. cycle ergometry), the observed occurrence of artefacts has previously been very low, at less than 1% (12, 20, 31). The criterion for the identification of artefacts also varies considerably among studies and, as such, they may also be significantly under-reported.

The increased occurrence of artefacts and the large bias and ES demonstrated that correction was necessary before HRM RR interval data could be considered NN (23). Only supine and standing recordings, with an artefact rate of <0.1% produced acceptable uncorrected HRV data. Except T1 and T6a, artefacts should be recognisable and correctable without the use of a simultaneous ECG recording. The potential for artefacts may be reduced by following the manufacturer's guidelines for fitting and positioning of the strap. External factors may also interfere with the transmission of the Bluetooth signal from the chest strap to the watch, particularly cross-body and wireless network interference (18). As the position of the watch was kept constant for all tests and participants and there were no other operating wireless devices in the laboratory, it is unlikely that these factors impacted the results of the present study.

The deletion of artefacts is the simplest of correction techniques, however as HRV spectral analysis is sensitive to changes in the signal length, discontinuity of the signal and loss of samples the current study does not support the use of deletion (29). Previous research that has used deletion to correct RR intervals should be treated with caution. Interpolation of artefacts maintains the length of the RR interval time series, overcoming the main

issue of deletion. Linear interpolation produced the smallest bias and ES in most, but not all cases. In line with the results of the present study, previous authors have found interpolation to increase the reliability of RR interval data (2). However, inserting linearly interpolated intervals can create flat sections with little or no variability, and it is advised that large sections of ectopic beats should not be edited using these techniques (23). From the results of the present study, it is not possible to state the number or frequency of artefacts that may be considered acceptable. However, caution should be given to the correction of RR intervals with sections that have more than four concurrent artefacts, or that contain a notable number within a small area of data. Future research could examine a threshold for the inclusion (or exclusion) of excessively noisy signals to determine whether to correct or discard the RR interval data.

The manual correction of RR intervals is time-consuming; Kubios HRV software provides options for the automatic detection and correction of ectopic beats and artefacts. However, despite Kubios appearing to accurately detect most artefacts, the results showed larger bias and effect sizes compared to both the ECG and the manually corrected cubic spline interpolation. Both correction methods used the same interpolation calculation in Matlab. Fig 2 shows a short section of data with a single type 4 artefact present; it is apparent that while Kubios software does accurately identify the artefact (blue), the erroneous beat is only replaced with a single interval, rather than the required two. While this does mean that the signal length remains constant, the number of intervals is one, or more, less than if correctly manually. It would appear that the reduced validity occurs because of this error. Until this issue has been addressed the automatic correction methods used within Kubios HRV software are not recommended for RR interval data recorded using HRM, where accuracy is important.
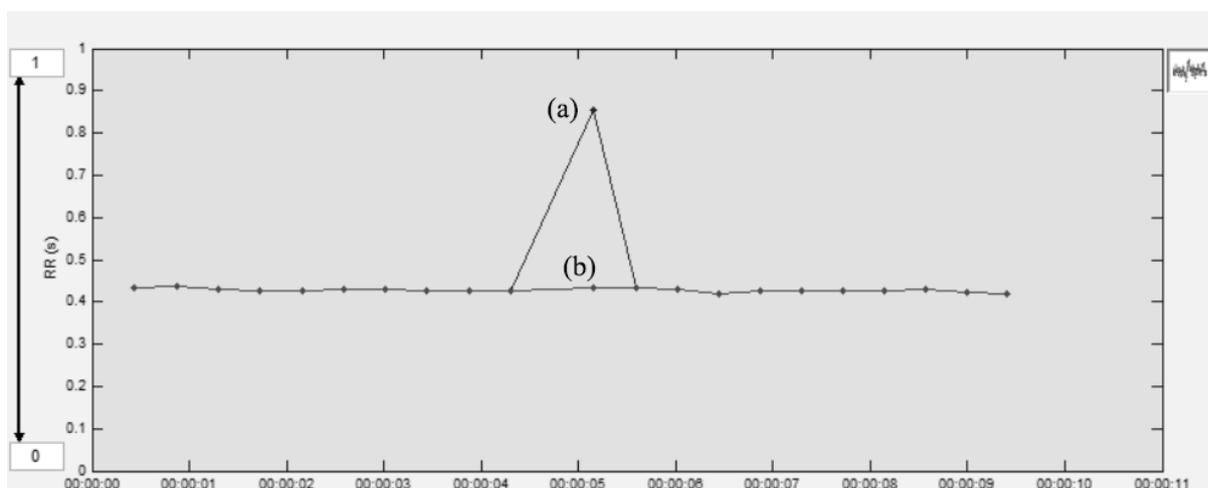

Fig 2: Kubios HRV software error correction, (a) original erroneous artefact (b) single cubic spline interpolated interval

PRACTICAL APPLICATIONS

Heart rate monitors are regularly used for the recording of RR intervals and the calculation of HRV in exercise related studies (12, 20, 31). However, both technical and biological artefacts occur in the RR interval data, which if left unedited, introduce bias and present a significant problem for the interpretation of calculated time, frequency and non-linear HRV parameters (23). The occurrence of artefacts in data requires correction in all but recordings with >0.1% of artefacts. The impact of exercise on the number of artefacts detected is likely to be related to the increase in heart rate and movement in the chest, that cause artefacts in the RR interval time series to increase relative to exercise intensity. As the results of the present study demonstrate, artefact correction is necessary in RR intervals obtained from HRMs. Correction of artefacts with a simple linear interpolation reduced bias and ES and increased ICC, in most but not all cases: caution should be given to RMSSD, LF:HF ratio, SD1 and SampEn at high (60+ $\dot{V}O_{2MAX}$) exercise intensities. Finally, the methods used in the present study are recommended for future studies using HRM in sport and exercise research, both at rest and during exercise.

GUIDELINES

Methods used for the correction of ectopic and erroneous signals vary considerably. It is recommended that researchers using HRM to record RR intervals state how the occurrence of artefacts was minimised, how they were identified and corrected. The following guidelines are proposed to increase transparency and minimise bias in RR interval recordings made using HRMs:

a) The fit and conductance of the chest strap should be maximised, by choosing a correctly sized strap appropriate for the participant, ensuring that the electrodes are wet, and the chest strap is stable. Minimise environmental sources of interference by checking the correct pairing of the HRM and the chest strap, and avoid other wireless networks. Position the watch (if not on the wrist) in front of the participant.

b) Where possible, select sections of RR intervals that are artefact free. If artefact free sections are not achievable, artefacts should be identified following the guidelines set out in the present paper. The percentage of artefacts identified should be stated in the methods, along with the inclusion/ exclusion criteria. Caution should be given to sections of data that contain more than four concurrent artefacts, or more than four artefacts in close proximity.

c) Identified artefacts must be corrected before HRV analysis. The means of artefact correction should be clearly stated. The findings of the present study support the use of linear interpolation.

ACKNOWLEDGEMENTS

REFERENCES

1.   Achten J and Jeukendrup AE. Heart rate monitoring. *Sports Med* 33: 517-538, 2003.

2.   Albrecht P and Cohen RJ. Estimation of heart rate power spectrum bands from real-world data: dealing with ectopic beats and noisy data. Presented at IEEE Computers in Cardiology Conference, Washington, D.C., 1988.

3.   Atkinson G and Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med* 26: 217-238, 1998.

4.   Barbosa R, Silva NT, Azevedo FM, Pastre CM, and Vanderlei LCM. Comparison of Polar® RS800G3™ heart rate monitor with Polar® S810i™ and electrocardiogram to obtain the series of RR intervals and analysis of heart rate variability at rest. *Clin Physiol Funct Imaging* 36: 112-117, 2014.

5.   Birkett C, Kienzle M, and Myers G. Interpolation over ectopic beats increases low frequency power in heart rate variability spectra. Presented at Computers in Cardiology Conference, Venice, Sep 23rd, 1991.

6.   Bland JM and Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 327: 307-310, 1986.

7.   Brennan M, Palaniswami M, and Kamen P. Do existing measures of Poincare plot geometry reflect nonlinear features of heart rate variability? *IEEE Trans Biomed Eng* 48: 1342-1347, 2001.

8.   Buchheit M, Solano R, and Millet GP. Heart-rate deflection point and the second heart-rate variability threshold during running exercise in trained boys. *Pediatr Exerc Sci* 19: 192-204, 2007.

9.   Citi L, Brown EN, and Barbieri R. A real-time automated point-process method for the detection and correction of erroneous and ectopic heartbeats. *IEEE Trans Biomed Eng* 59: 2828-2837, 2012.

10.   Cohen J. *Statistical power analysis for the behavioral sciences.* Oxford, UK: Academic press, 2013.

11.   de Lima DC, Ribeiro HS, Cristina R, Oliveira M, de Vasconcelos Generoso S, Lima AS, and Correia MITD. Functional status and heart rate variability in end-stage liver disease patients: association with nutritional status. *Nutrition* 31: 971-974, 2015.

12.   De Vito G, Galloway S, Nimmo MA, Maas P, and McMurray JJ. Effects of central sympathetic inhibition on heart rate variability during steady- state exercise in healthy humans. *Clin Physiol Funct Imaging* 22: 32-38, 2002.

13.   Duarte TS, Júnior AJF, Rezende EC, Pertence LC, Carlos J, Marins B, and de Lima JRP. Association between heart rate variability threshold and post-exercise vagal reactivation. *J Exerc Physiol Online* 17: 58-65, 2014.

14.   Fronchetti L, Nakamura FY, De-Oliveira FR, Lima-Silva AE, and Lima J. Effects of high-intensity interval training on heart rate variability during exercise. *J Exerc Physiol Online* 10: 1-9, 2007.

15.   Gamelin F, Baquet G, Berthoin S, and Bosquet L. Validity of the polar S810 to measure RR intervals in children. *Int J Sports Med* 29: 134-138, 2008.

16.   Gamelin FX, Berthoin S, and Bosquet L. Validity of the polar S810 heart rate monitor to measure RR intervals at rest. *Med Sci Sports Exerc* 38: 887-893, 2006.

17.   Giles D, Draper N, and Neil W. Validation of the Polar V800 heart rate monitor to measure RR intervals at rest. *Eur J Appl Physiol* 116: 563-571, 2016.

18.   Kara A and Bertoni HL. Blockage/shadowing and polarization measurements at 2.45 ghz for interference evaluation between Bluetooth and IEEE 802.11 WLAN. Presented at IEEE Antennas and Propagation Society International Symposium, 2001, 2001.

19.   Karim N, Hasan JA, and Ali SS. Heart rate variability–a review. *J Basic Appl Sci* 7: 71-77, 2011.

20.   Lewis MJ, Kingsley M, Short A, and Simpson K. Influence of high-frequency bandwidth on heart rate variability analysis during physical exercise. *Biomed Signal Process Control* 2: 34-39, 2007.

21.   Mateo M, Blasco-Lafarga C, Martínez-Navarro I, Guzmán JF, and Zabala M. Heart rate variability and pre-competitive anxiety in BMX discipline. *Eur J Appl Physiol* 112: 113-123, 2012.

22.   Morales J, Garcia V, García-Massó X, Salvá P, and Escobar R. The use of heart rate variability in assessing precompetitive stress in high-standard judo athletes. *Int J Sports Med* 34: 144-151, 2013.

23.   Peltola MA. Role of editing of R–R intervals in the analysis of heart rate variability. *Front Physiol* 3, 2012.

24.   Pöyhönen M, Syväoja S, Hartikainen J, Ruokonen E, and Takala J. The effect of carbon dioxide, respiratory rate and tidal volume on human heart rate variability. *Acta Anaesthesiol Scand* 48: 93-101, 2004.

25.   Quintana DS, Heathers JA, and Kemp AH. On the validity of using the Polar RS800 heart rate monitor for heart rate variability research. *Eur J Appl Physiol* 112: 4179-4180, 2012.

26.   Salo MA, Huikuri HV, and Seppanen T. Ectopic beats in heart rate variability analysis: effects of editing on time and frequency domain measures. *Ann Noninvasive Electrocardiol* 6: 5-17, 2001.

27.   Sami S, Seppänen M, and Kuusela A. Artefact correction for heart beat interval data. Presented at 1st Advanced Methods for Processing Bioelectrical Signals Meeting, Jyväskylä, Finland: University of Jyväskylä, Oct 1st, 2004.

28.   Tarvainen MP, Niskanen J-P, Lipponen JA, Ranta-Aho PO, and Karjalainen PA. Kubios HRV–heart rate variability analysis software. *Comput Methods Programs Biomed* 113: 210-220, 2014.

29.   Task. Force of the European Society of Cardiology & the North American Society of Pacing and Electrophysiology. Standards of measurement, physiological interpretation, and clinical use. *Circulation* 93: 1043-1065, 1996.

30.   Thomas JR, Nelson JK, and Silverman SJ. *Research methods in physical activity.* Champaign, IL: Human Kinetics, 2010.

31.   Tulppo MP, Hughson RL, Mäkikallio TH, Airaksinen KJ, Seppänen T, and Huikuri HV. Effects of exercise and passive head-up tilt on fractal and complexity properties of heart rate dynamics. *Am J Physiol* 280: H1081-H1087, 2001.

32.   Vanderlei L, Silva R, Pastre C, Azevedo FM, and Godoy M. Comparison of the Polar S810i monitor and the ECG for the analysis of heart rate variability in the time and frequency domains. *Braz J Med Biol Res* 41: 854-859, 2008.

33.   Wallén MB, Hasson D, Theorell T, Canlon B, and Osika W. Possibilities and limitations of the polar RS800 in measuring heart rate variability at rest. *Eur J Appl Physiol* 112: 1153-1165, 2012.

34. Weippert M, Behrens M, Rieger A, and Behrens K. Sample Entropy and Traditional Measures of Heart Rate Dynamics Reveal Different Modes of Cardiovascular Control During Low Intensity Exercise. *Entropy* 16: 5698-5711, 2014.

35. Weippert M, Kumar M, Kreuzfeld S, Arndt D, Rieger A, and Stoll R. Comparison of three mobile devices for measuring R–R intervals and heart rate variability: Polar S810i, Suunto t6 and an ambulatory ECG system. *Eur J Appl Physiol* 109: 779-786, 2010.

36. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 19: 231-240, 2005.

PRE-PRINT

PRE-PRINT