Repeatability and Reliability of Heart Rate Variability in Healthy, Adult Pony Mares

Elize van Vollenhoven[a*], Catharina Cornelia Grant[b], Lizelle Fletcher[c], André Ganswindt[d],

Patrick Collin Page[a]

[a] Department of Companion Animal Clinical Studies, Faculty of Veterinary Science, University of

Pretoria, Private Bag X04, Onderstepoort, 0110, South Africa, elize.vanvollenhoven@up.ac.za;

patrick.page@up.ac.za

[b] Section Sports Medicine, Faculty of Health Sciences, University of Pretoria, P.O. Box 37897, Fairie

Glen, 0043,South Africa,

[c] Department of Statistics, University of Pretoria, Private Bag X20, Hatfield, 0028 South Africa;

lizelle.fletcher@up.ac.za

[d] Endocrine Research Laboratory, Department of Anatomy and Physiology, Faculty of Veterinary

Science, University of Pretoria, P.O. Box X04, Onderstepoort, 0110, South Africa;

andre.ganswindt@up.ac.za

*Corresponding author: Dr Elize van Vollenhoven, Department of Companion Animal Clinical

Studies, Faculty of Veterinary Science, University of Pretoria, Ou Soutpan Rd M35, Onderstepoort,

Gauteng, South Africa.

E-mail address: elize.vanvollenhoven@up.ac.za (E. van Vollenhoven)

Highlights

- Choice of software correction factor (CF) influences HRV indicators in horses.
- HRV indicators have good repeatability but overall poor-to-good reliability.

- HRV indicators obtained from herringbone stocks appear less reliable than pasture.

- Using normalized LF and HF improved reliability for frequency domain indicators.

- The CF used should be defined and defendable to enhance standardization of HRV.

Abstract

Heart rate variability (HRV) is an important non-invasive method to quantify stress by measuring sympathetic and parasympathetic activity of the autonomic nervous system. Few studies exist on the repeatability and reliability of HRV in equids. The objectives of this study were to (a) compare the effect of different correction factors (CF) available in HRV analysis software on HRV indicator values and (b) to determine the repeatability and reliability of HRV indicators in an unrestricted (pasture) and a restricted movement (equine examination stocks) environment. Data were recorded on five consecutive days from six adult pony mares. Short term tachograms were compared with regards to software CF by graphical and statistical (Friedman's and Wilcoxon signed rank test) comparison. The results showed that the specific CF influences the HRV indicator values. The Strong CF was able to balance the elimination of artefacts without removing the variability of RR-intervals and was subsequently used to determine repeatability and reliability. HRV indicators showed good repeatability over the 5 days using Friedman's test (pasture: $p=0.162$-0.898; examination stocks: $p=0.29$-0.865), indicating that there were no significant differences between HRV indicator values. The reliability, represented by intraclass correlation coefficient (ICC) and coefficient of variation (CV), was poor to good for pasture data (ICC=0.44-0.79; CV=10-68.10) and examination stocks data (ICC=0.22-0.83; CV=16.07-62.29). Measurements obtained from the examination stocks were less reliable than those from pasture. Using normalized low-frequency and normalized high-frequency components improved reliability. Free-movement environment based HRV recordings could ensure better reliability, but may require the use of a stronger CF.

1.   Introduction

Stress in humans and animals can be assessed by invasive methods which include collection of blood samples for measurement of stress-related hormones, metabolic and physiological parameters [1-3] and arterial catheterization for direct measurement of blood pressure [2]. The disadvantage of using invasive methods is that they can influence the stress the individual is experiencing, thus not providing an accurate estimate of the actual stress experienced. While non-invasive methods can provide a more accurate indication of the stress experienced by the animal, the method selected needs to be both valid and quantifiable. Non- or minimally invasive methods include recording behavioral indices [4], assessment of salivary cortisol levels [5], fecal or urinary glucocorticoid metabolite levels [5, 6], indirect blood pressure measurements [4], heart rate monitoring [7], and quantification of heart rate variability (HRV) [5, 6].

Heart rate variability, referring to the changes in beat to beat heart rate measured over a period of a RR-interval recording (tachogram) [8, 9], has been proven as a valid method to assess stress in humans [8] and animals [10], including horses [5, 6, 11-14], pigs [15, 16], sheep and goats [17-19], cattle [20, 21], poultry [22], and dogs [23, 24]. Studies in humans showed that standardization of methodology is important to assure inter- and intra-study repeatability of HRV measurements [25-29]. These studies indicated that inconsistent results can be expected with data sampling at different times of the day [25], presence of heart disorders or unhealthy participants [25, 28], gender differences [26], or when different HRV indicators are used for monitoring the autonomic system [26].

Using the correct methodology, HRV quantification can be applied as a non-invasive indicator of autonomic control, which is invaluable during non-verbal stress and/or pain evaluation in humans and especially in animals. However, the general availability of automatic RR-detection systems and software makes it easy to underestimate the complexity of the technique and the need for species-specific standardization. Although widely used in equine research, little is known on how appropriate it is to use this technique in horses and what the methodological pitfalls are. During the last decade HRV quantification methods and techniques in humans were applied directly in animal studies with low comparability between studies [30]. Specifically, in equids several confounding factors exist which may influence the repeatability and reliability of HRV quantification. For example, it is

difficult to accurately determine HRV in horses from automatically detected RR-intervals, due to the prominent T-wave which may be misinterpreted as an R-peak [30, 31]. This implies that two short inter-beat-intervals are identified instead of one [30]. Methodological issues reported by Stucke et al [30] also highlighted the movement of horses while measuring RR-intervals and the fact that RR-interval series recorded should be longer than the actual 5-minute tachogram that is used for HRV quantification. This longer interval series is recommended due to a relatively high incidence of first and second degree AV-blocks as well as other arrhythmias in resting horses in response to dominant vagal activity [32]. Automatic assumption of the applicability of methodology used in human or other animal studies during HRV quantification in horses may provide an explanation for the low comparability between studies. In HRV studies it is often not reported in the methodology if a correction factor (CF) was used during data analysis and if reported no reason for the specific choice is indicated [33-36].

Limited information exists about the effect of using the various CF available in HRV analysis software to automatically correct artefacts [6, 37] as well as the species-specific repeatability and reliability of HRV quantification, especially for horses. Repeatability depicts the variation in replicated measurements in the same animal under equivalent situations. Relative reliability is the variability of the different measurements of the same animal relative to the total variation of all the animals in a study, as represented by intraclass correlation coefficient (ICC) [38-41]. Absolute reliability depicts the variability of repeated measurements relative to the mean on the same animal, as represented by the coefficient of variation (CV) [39].

The aims of the present study were, firstly to determine the effect of using different CF available in HRV analysis software (i.e. repeatability) on HRV indicator values, and secondly to determine the repeatability and reliability of time-domain, frequency-domain and geometric (Poincaré plot) short-term indicator values measured in healthy, adult pony mares.

2. Material and Methods

2.1. Study animals

Seven healthy, adult, non-pregnant Nooitgedacht pony mares, (mean ± SD) age 9.5 ± 4.8 years, mass 415 ± 26 kg from the Onderstepoort Teaching Animal Unit were randomly selected for the study. The data sets available from six of the mares were analyzed as one mare's data set was omitted due to incomplete data obtained. Mares were fed ad libitum *Eragrostis curvula* hay at maintenance requirement level and had free access to water, except during the monitoring phase in the examination stocks. The mares were kept in their normal pasture habitat (~ 1 Ha) between data collection. Only mares determined clinically healthy (based on physical examination, including comprehensive cardiac auscultation, conducted within six days of HRV data collection), with normal habitus and appetite observed on the morning of commencement of the study, and that were not utilized in any other research program during the 30 days prior to the start of data collection, were included. The study was conducted in winter (Southern hemisphere), when the mares were in anestrus (confirmed by rectal palpation records), to standardize reproductive status [10].

This study was approved by the Animal Ethics Committee of the University of Pretoria (Study no. V034-13) and no animal welfare concerns were observed.

2.2. Experimental procedures

The mares were familiar with the researcher, the study environment (equine herringbone examination stocks and pasture), and were already accustomed to wearing the HRV recording equipment. The procedures described below were followed on a daily basis on five consecutive days. The portable heart rate monitors (Polar® RS800, Polar® Electro Öy, Kempele, Finland) and the heart rate monitor belts (WearLink belts, Polar® Electro Öy, Kempele, Finland) were attached by the primary researcher to the individual mares from 8h00, while grouped in the same pasture. The heart rate monitor and the heart rate monitor belt were attached to the mare's thorax by means of a surcingle. The heart rate monitor belt (containing the electrodes and transmitter pocket) transmitted data to the heart rate monitor. The transmitter pocket was placed mid-left thorax (as prescribed by the manufacturer) and physically adjusted to eliminate artefacts due to pronounced T-waves i.e. elevated

heart rate displayed for a horse at rest [10]. To promote signal transmission ECG gel was applied to the electrode site, which had been clipped not less than six days prior to data collection, and cleaned with alcohol. RR-intervals were recorded on the pasture for (mean ± SD) 113 ± 27-minutes to obtain short term tachograms [8, 10] that represented the pasture environment (unrestricted movement). During data recording the mares were observed from a distance and only approached by the researcher to check on the heart rate monitor functionality every 15 minutes, but were not handled unless the monitor belt had to be adjusted. These time checks and adjustments were recorded. After the HRV data (RR-intervals) were recorded in the pasture the mares were walked in hand to the stocks (approximately 11h00) and placed in adjacent individual stocks. RR-intervals were recorded in the stocks for 76 ± 7-minutes (until approximately 13h00) to obtain short term tachograms [8, 10] that represented the stocks environment (restricted movement). The surcingles, heart rate monitors and belts were removed following data collection in the stocks and the mares were returned as a group to their normal pasture, where they remained overnight. The data obtained with the heart rate monitors were downloaded daily to a computer for analysis.

2.3. Environmental data

Ambient temperature was recorded by data loggers (iButton$^®$ DS1923 and Coldchain Thermo Dynamics Software, Fairbridge Technologies CC, Wendywood, South Africa) placed in the camps housing the mares and in the stocks. Ambient temperature during the HRV recording time ranged daily from 14-17$^°$C (mean 15$^°$C) on the pastures and 9-22$^°$C (mean 18$^°$C) in the stocks. No rain was recorded during the data collection period.

3. Data processing and analysis

Data were downloaded from the monitor using Polar$^®$ ProTrainer 5 (Polar$^®$ Electro Europe BV, Fleurier Branch, Switzerland) adapted for horses and then transferred to the HRV Analysis Software 1.1. for Windows or Kubios (The Biomedical Signal Analysis Group, Department of Applied Physics, University of Kuopio, Finland), where the variability of the RR-intervals were quantified with the aid

of time domain, frequency domain and Poincaré plot analyses. Low-frequency (LF) and high-frequency (HF) bands were set at 0.01-0.07 Hz and 0.07-0.6 Hz, respectively [34, 42-44].

As has been suggested [30] the RR-interval series sampled on the pasture and stocks were longer than the actual standardized 5-minute tachograms [8, 10] necessary for short-term HRV quantification. R-wave errors were then eliminated by visual inspection of tachograms and selection of the 5-minute section with the least number of artefacts as an accepted methodology to minimize errors [31, 34, 45, 46], followed by automatic correction of artefacts with the aid of mathematical algorithms (Kubios software). The detrending procedure was based on smoothness priors set at 500 ms as described by Tarvainen et al [47]. Correction filters were set at Low, Medium, Strong and Very Strong which identified RR-intervals, differing respectively with 0.45, 0.25, 0.15 and 0.05 s from the local mean RR-interval, as artefacts. The software then replaced these artefacts with interpolated intervals (computed from the difference between the previous and next approved RR-intervals) [48].

Heart rate measures, namely MeanRR = mean R-R interval (inter-beat interval or time interval between two consecutive heart beats measured in ms) and MeanHR = mean heart rate (bpm), were measured as well as the following HRV indicators:

- Time domain indicators: SDNN = standard deviation of normal-to-normal intervals (ms), RMSSD = root mean squared differences of the standard deviation (ms), pNN50 = percentage of beats that changed more than 50 ms from the previous beat (%);

- Frequency domain indicators: LF = Low frequency power obtained with auto-regressive spectral analysis of RR-intervals (ms^2), HF = power obtained with auto-regressive spectral analysis of RR-intervals (ms^2), LF/HF = low frequency to high frequency ratio; LF nu = low frequency power normalized units ($\frac{LF}{total\ power-VLF}$), HF nu = high frequency power normalized units $\left(\frac{HF}{total\ power-VLF}\right)$, (VLF = very low frequency);

- The Poincaré plot (graphical representation of the RR-interval plotted against the previous RR-interval) [6] : SD1 = standard deviation 1 derived from Poincaré plot (ms), SD2 = standard deviation 2 derived from Poincaré plot (ms).

3.1. Statistical analysis

Statistical analysis was performed using SPSS® Statistics version 22 for Windows (IBM Corp, Armonk NY, USA). The significance level was set at 0.05.

3.1.1. The influence of different correction factors on heart rate measures and HRV indicator values

The repeatability of the different CF, i.e. the differences between the measurements per HRV indicator in the same horse, sampled on five consecutive days under equivalent conditions, was determined using the non-parametric Friedman's test. HRV values were also graphically compared as shown in Fig. 1 - 3. When significant results were obtained Wilcoxon signed rank tests were performed *post hoc* to determine the statistical differences between the various CF. Kendall's Coefficient of Concordance (W), as an indication of effect size, was also calculated.
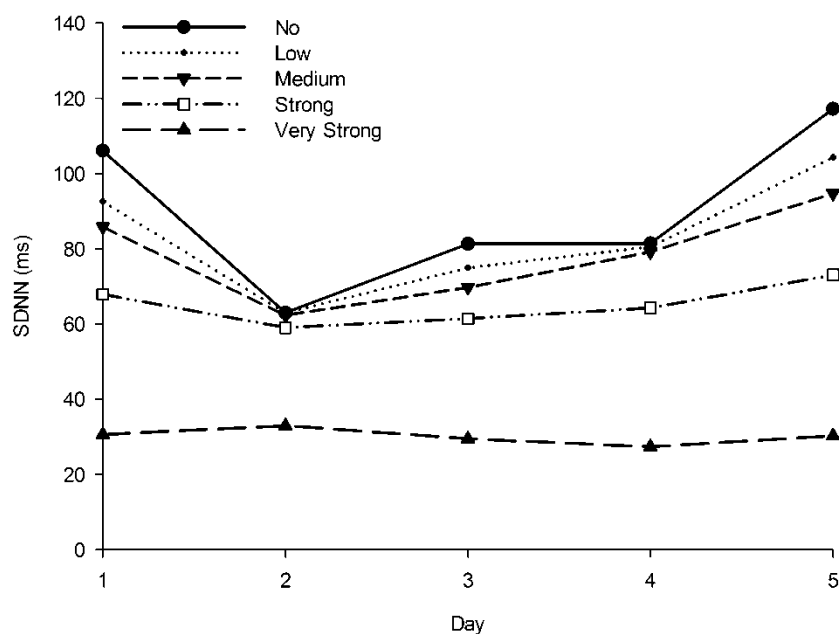

Fig. 1. Graphical representation of the mean values of the different correction factors for SDNN (pasture environment) compared on 5 consecutive days.
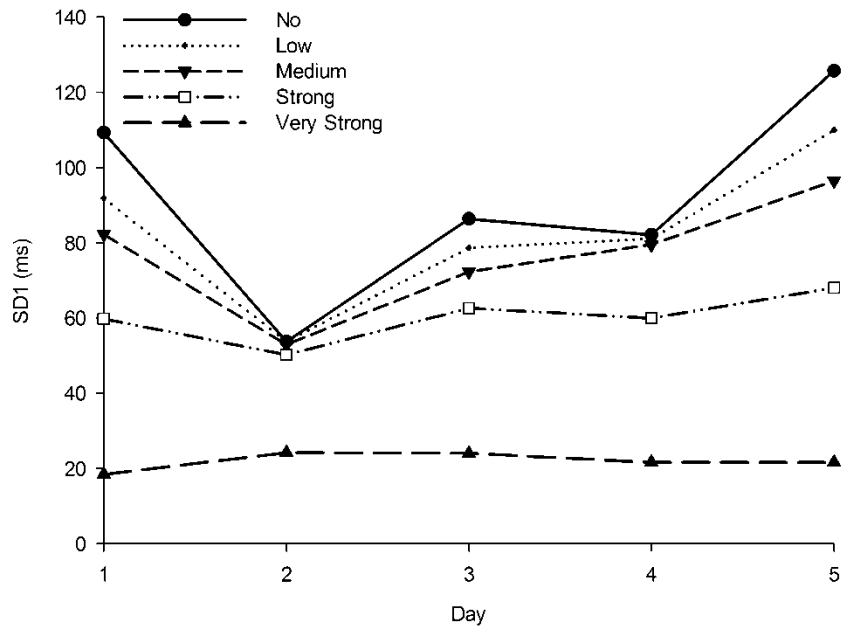
Fig. 2. Graphical representation of the mean values of the different correction factors for SD1 (pasture environment) compared on 5 consecutive days.
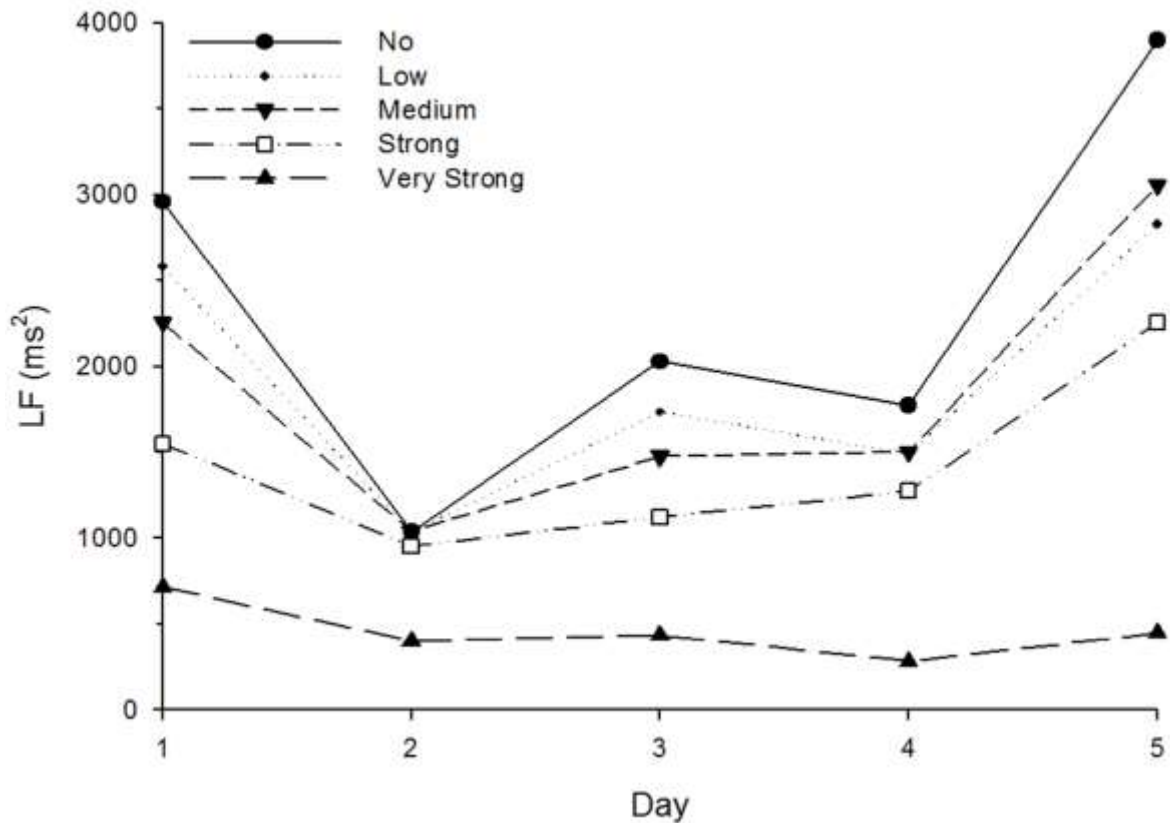
Fig. 3. Graphical representation of the mean values of the different correction factors for LF (pasture environment) compared on 5 consecutive days.

3.1.2. Repeatability of heart rate measures and HRV indicators using CF Strong

The CF choice for the rest of the data analysis was determined by two factors. In the first instance it is critically important to use a data set that is cleared from artefacts and background noise for HRV quantification. However, it is just as important not to remove all variation with a too strong correction factor as this will result in removal of the variability of the RR-intervals [8]. Thus, an appropriate CF (Strong) as determined by graphical and statistical comparisons was then selected and applied to determine if there were differences in HRV from tachograms obtained on the five consecutive days.

Repeatability of HRV indicators was determined with Friedman's test over the five days and Wilcoxon signed rank tests on the significant results to determine which days specifically differ from each other.

3.1.3. Reliability of heart rate measures and HRV indicators using CF Strong

Reliability of the HRV indicators was assessed by means of ICC and the CV [29, 39]. The CV was

calculated as the average of the individual CV's for each horse where $CV = \frac{sd}{\bar{x}} \times 100\%$ [39]. The ICC

was calculated by SPSS® Statistics software using a 2-way mixed model with measures of

consistency.

4. Results

A comparison between CF in both the pasture (Table 1; Fig. 1- 3) and stocks environment (Table

2) indicated that there were no significant changes between the HRV indicator values and heart rate

measures when comparing No, Low and Medium with each other, except for LF normalized (No vs.

Medium in the pasture and in the stocks). However, between the Strong and Very Strong CF there

were significant differences for the majority of HRV indicators (pasture: 10/10; stocks: 7/10) and for

heart rate measures (pasture: 2/2; stocks: 0/2). The Very Strong CF was not considered applicable as

it consistently differed from the other four factors. When comparing the Strong and Very Strong CF

with No, Low and Medium CF, respectively, the heart rate measures (pasture: 5/12; stocks: 6/12) and

the HRV indicators comparisons (pasture: 50/60; stocks: 32/60) showed significant differences (Table

1 and 2).

A similar pattern was demonstrated for HRV indicator values obtained with the No, Low, Medium

and Strong CF, illustrated graphically by SDNN, SD1 and LF (Fig. 1, 2 and 3). It was therefore

decided to continue with the Strong CF to evaluate repeatability. The Kendall's Coefficient of

Concordance, calculated for Friedman's test, ranged from 0.050 to 0.272.

Table 1
Statistical results: comparison between combinations of correction factors (Kubios) over five days applied to heart rate measures and heart rate variability indicators from six pony mares in a pasture environment

Heart rate measures	Friedman test (p-value)	Wilcoxon signed rank test (p-value)									
		No vs. Low	No vs. Medium	No vs. Strong	No vs. Very Strong	Low vs. Medium	Low vs. Strong	Low vs. Very Strong	Medium vs. Strong	Medium vs. Very Strong	Strong vs. Very Strong
MeanRR	<0.001**	0.125	0.063	0.031*	0.844	0.063	0.031*	0.563	0.156	0.094	0.031*
MeanHR	0.001**	0.875	0.625	0.031*	0.031*	0.063	0.156	0.063	0.688	0.031*	0.031*
HRV Indicators											
SDNN	<0.001**	0.125	0.063	0.031*	0.031*	0.063	0.031*	0.031*	0.031*	0.031*	0.031*
RMSSD	<0.001**	0.125	0.063	0.031*	0.031*	0.063	0.031*	0.031*	0.031*	0.031*	0.031*
PNN50	<0.001**	0.250	0.063	0.031*	0.031*	0.063	0.031*	0.031*	0.031*	0.031*	0.031*
LF	<0.001**	0.125	0.063	0.031*	0.031*	0.625	0.031*	0.031*	0.094	0.031*	0.031*
HF	<0.001**	0.125	0.063	0.031*	0.031*	0.063	0.031*	0.031*	0.031*	0.031*	0.031*
LF/HF	<0.001**	0.125	0.063	0.156	0.031*	0.063	0.313	0.031*	0.563	0.031*	0.031*
LF nu	<0.001**	0.125	0.031*	0.156	0.031*	0.063	0.219	0.031*	0.563	0.031*	0.031*
HF nu	0.001**	0.125	0.125	0.219	0.031*	0.188	0.563	0.031*	0.688	0.031*	0.031*
SD1	<0.001**	0.125	0.063	0.031*	0.031*	0.063	0.031*	0.031*	0.031*	0.031*	0.031*
SD2	<0.001**	0.125	0.063	0.031*	0.031*	0.063	0.031*	0.031*	0.031*	0.031*	0.031*

*$P<0.05$, **$P<0.01$.

HRV = heart rate variability; RR = RR-interval; HR = heart rate; SDNN = standard deviation of RR-interval; RMSSD = root mean square of successive differences in RR-intervals; pNN50 = percentage of intervals differing by >50 ms (milliseconds) from preceding interval; HF = high-frequency components; LF = low-frequency components; LF/HF = autonomic balance;; LF nu = low frequency power normalized units; HF nu = high frequency power normalized units; SD1 = standard deviation of short term variability; SD2 = standard deviation of the long-term variability.

Table 2
Statistical results: comparison between combinations of correction factors (Kubios) over five days applied to heart rate measures and heart rate variability indicators from six pony mares in an equine herringbone examination stocks environment

Wilcoxon signed rank test (p-value)

Heart rate measures	Friedman test (p-value)	No vs. Low	No vs. Medium	No vs. Strong	No vs. Very Strong	Low vs. Medium	Low vs. Strong	Low vs. Very Strong	Medium vs. Strong	Medium vs. Very Strong	Strong vs. Very Strong
MeanRR	0.039*	0.875	0.125	1	0.031*	0.125	0.438	1.000	0.688	0.031*	0.156
MeanHR	0.002**	0.250	0.250	0.031*	0.031*	0.250	0.688	0.031*	0.156	0.031*	0.063
HRV Indicators											
SDNN	<0.001**	0.125	0.125	0.031*	0.031*	0.125	0.219	0.031*	0.438	0.031*	0.031*
RMSSD	<0.001**	0.125	0.125	0.031*	0.031*	0.125	0.313	0.031*	0.438	0.031*	0.031*
PNN50	<0.001**	0.125	0.250	0.031*	0.031*	0.250	0.031*	0.031*	0.031*	0.031*	0.031*
LF	<0.001**	0.125	0.125	0.031*	0.031*	0.250	0.031*	0.031*	0.219	0.031*	0.031*
HF	<0.001**	0.125	0.125	0.031*	0.031*	0.125	0.438	0.031*	0.438	0.031*	0.031*
LF/HF	0.029*	0.125	0.125	0.063	0.063	0.125	0.313	0.063	0.438	0.063	0.063
LF nu	0.060	0.250	0.031*	0.063	0.063	0.625	0.438	0.063	0.563	0.063	0.094
HF nu	0.005**	0.125	0.125	0.031*	0.063	0.250	0.219	0.063	0.219	0.063	0.094
SD1	<0.001**	0.125	0.125	0.031*	0.031*	0.125	0.313	0.031*	0.438	0.031*	0.031*
SD2	<0.001**	0.125	0.125	0.031*	0.031*	0.250	0.219	0.031*	0.438	0.031*	0.031*

*$P<0.05$, **$P<0.01$.

HRV = heart rate variability; RR = RR-interval; ms = millisecond; HR = heart rate; bpm = beats per minute; SDNN = standard deviation of RR-interval; RMSSD = root mean square of successive differences in RR-intervals; pNN50 = percentage of intervals differing by >50 ms from preceding interval; LF = low-frequency components; HF = high-frequency components; LF/HF = autonomic balance; LF nu = low frequency power normalized units; HF nu = high frequency power normalized units; SD1 = standard deviation of short term variability; SD2 = standard deviation of the long-term variability; SD = Standard Deviation, CV = coefficient of variation; ICC = intraclass correlation coefficient.

Tables 3 and 4 show the mean, standard deviation and median of the HRV indicators as well as the measures of reliability (CV and ICC) and the Friedman test p-values, measured on five consecutive days during the pasture and stocks monitoring periods.

Comparing the different HRV indicator values and heart rate measures on five separate days in the pasture (Table 3) and stocks (Table 4), revealed that there were no significant differences between the respective HRV indicators in either of the two environments. However, the heart rate measures (MeanHR and MeanRR) differed significantly in both these environments. From the *post hoc* tests it was evident that only Day 2 differed significantly from the other days for heart rate measurements on pasture (p=0.016 for MeanHR and MeanRR for Day 1 vs. Day 2) as well as in the stocks (p=0.031 for MeanHR and MeanRR Day 1 vs. Day 2; p=0.016 for MeanRR and MeanHR for Day 2 vs. Day 3 and Day 2 vs. Day 5; p=0.047 for MeanRR and p=0.109 for MeanHR for Day 4 vs. Day 5).

The CV and ICC for heart rate measures and HRV indicators obtained from pastures (Table 3) varied between 10.00 – 68.10 and 0.44 – 0.79, respectively. These values for the stocks environment (Table 4) varied between 8.78 – 62.29 and 0.22 – 0.95, respectively.

Table 3. Heart rate measures and heart rate variability: mean ± SD (median) values and statistical results from six pony mares in a pasture environment (Strong correction factor).

Heart rate measures	Day 1	Day 2	Day 3	Day 4	Day 5	p-value	CV	ICC
MeanRR (ms)	1998.24 ± 218.85 (1929.42)	1645.59 ± 225.83 (1538.31)	2279.94 ± 246.95 (2281.70)	1893.46 ± 295.72 (1950.32)	2074 ± 456.16 (1897.01)	0.007**	15.31	0.74
MeanHR (bpm)	30.48 ± 3.10 (31.28)	37.08 ± 4.59 (39.08)	26.69 ± 3.05 (26.36)	32.42 ± 5.24 (30.81)	30.09 ± 5.69 (31.93)	0.007**	15.61	0.76
HRV Indicators								
SDNN (ms)	66.42 ± 25.19 (74.24)	60.22 ± 12.28 (61.72)	61.41 ± 18.76 (61.17)	60.51 ± 26.54 (58.71)	72.93 ± 18.43 (75.80)	0.316	24.49	0.76
RMSSD (ms)	78.20 ± 33.27 (79.92)	70.52 ± 11.57 (69.19)	88.12 ± 25.10 (89.49)	77.49 ± 35.41 (67.77)	95.13 ± 16.58 (92.10)	0.162	26.91	0.59
pNN50 (%)	40.25 ± 21.42 (40.57)	46.43 ± 8.61 (47.59)	49.96 ± 14.52 (53.24)	42.70 ± 16.77 (41.41)	57.74 ± 9.25 (55.95)	0.419	28.52	0.44
LF (ms^2)	1550.33±1167.44 (1441.81)	953.54±403.23 (1074.63)	1122.45±1027.39 (693.96)	1272.38±1148.72 (1069.33)	2254±2369.64 (1831.33)	0.898	68.10	0.65
HF (ms^2)	2998.07±1915.57 (3423.25)	2690.10±1264.37 (2803.30)	3069.87±1884.62 (2576.90)	3250.27±2608.96 (2638.66)	3884.46±1305.83 (3823.77)	0.450	43.09	0.68
LF/HF	0.52 ± 0.30	0.39 ± 0.19	0.32 ± 0.16	0.37 ± 0.16	0.50 ± 0.47	0.510	45.90	0.78
LF nu	37.48±18.20 (39.07)	29.78±9.94 (29.02)	25.67±10.66 (24.60)	27.9±8.29 (25.05)	31.79±21.74 (32.12)	0.510	37.38	0.78
HF nu	76.37±9.81 (75.85)	79.26±8.96 (80.07)	82.68±6.96 (83.37)	79.45±8.82 (83.70)	77.92±18.27 (79.43)	0.623	10.00	0.79
SD1 (ms)	55.49±23.62 (56.69)	50±8.21 (49.05)	62.55±17.82 (63.52)	54.98±25.15 (48.07)	67.51±11.81 (65.32)	0.162	26.92	0.59
SD2 (ms)	75±28.98 (83.85)	68.68±16.82 (69.45)	58.91±21.68 (51.52)	65.02±28.94 (62.37)	76.54±26.74 (85.49)	0.450	27.33	0.77

*P<0.05, **P<0.01

HRV = heart rate variability; SD = Standard Deviation,

RR = RR-interval; ms = millisecond; HR = heart rate; bpm = beats per minute; SDNN = standard deviation of RR-interval; RMSSD = root mean square of successive differences in RR-intervals; pNN50 = percentage of intervals differing by >50 ms from preceding interval; SD1 = standard deviation of short term variability; SD2 = standard deviation of the long-term variability; HF = high-frequency components; LF = low-frequency components; LF/HF = autonomic balance; LF nu = low frequency power normalized units; HF nu = high frequency power normalized units;

CV = coefficient of variation; ICC = intraclass correlation coefficient.

Table 4. Heart rate measures and heart rate variability: mean ± SD (median) values and statistical results from six pony mares in an equine herringbone examination stocks environment (Strong correction factor).

Heart rate measures	Day 1	Day 2	Day 3	Day 4	Day 5	p-value	CV	ICC
MeanRR (ms)	2099.54±329.06 (1979.49)	1901.29±283.35 (1877.16)	2289.92±366.51 (2345.2)	2104.71±344.64 (2178.56)	2210.03±296.28 (2253.34)	0.001**	8.78	0.95
MeanHR (bpm)	30.89±7.18 (30.42)	32.35±4.75 (32.10)	27.06±4.80 (25.66)	29.52±5.33 (27.77)	28.03±3.93 (27.26)	0.001**	11.44	0.84
HRV Indicators								
SDNN (ms)	101.64±130.03 (57.94)	51.56±17.51 (54.66)	58.01±24.81 (50.62)	59.29±26.80 (62.65)	64.14±23.03 (72.82)	0.450	32.83	0.48
RMSSD (ms)	134.17±175.25 (70.31)	57.11±16.27 (56.89)	73.83±27.62 (71.70)	72.61±28.37 (74.44)	68.55±19.00 (66.52)	0.676	35.32	0.32
pNN50 (%)	42.86±18.83 (46.51)	30.65±15.34 (29.52)	44.38±17.95 (43.79)	41.64±20.08 (40.68)	38.96±14.66 (39.25)	0.377	30.47	0.83
LF (ms^2)	3029.62±4942.24 (1071.03)	1219.56±970.11 (1238.38)	1675.93±1689.21 (963.65)	1687.25±1348.04 (1880.21)	2696.66±2392.45 (2822.66)	0.865	62.11	0.67
HF (ms^2)	30102.71±69644.23 (1550.26)	1633.08±1054.06 (1593.43)	2193.98±1510.79 (1606.21)	1994.3±1242.18 (2153.86)	5529.96±7559.03 (2360.46)	0.587	62.29	0.22
LF/HF	0.49±0.45 (0.45)	0.78±0.38 (0.73)	0.63±0.33 (0.69)	0.71±0.38 (0.73)	0.66±0.51 (0.56)	0.675	53.5	0.70
LF nu	33.95±24.77 (35.93)	47.99±14.46 (48.82)	42.68±18.46 (46.89)	44.81±18.21 (49.3)	39.96±22.30 (41.16)	0.730	40.57	0.74
HF nu	79.5±14.86 (79.72)	67.23±13.77 (67.92)	72.39±11.87 (68.92)	69.11±14.63 (67.32)	73.00±18.51 (73.59)	0.290	16.07	0.67
SD1 (ms)	95.18±124.28 (49.93)	40.52±11.56 (40.35)	52.42±19.62 (50.92)	51.54±20.16 (52.82)	48.66±13.51 (47.23)	0.675	35.32	0.32
SD2 (ms)	107.35±136.77 (62.6)	60±22.26 (65.24)	62.03±31.60 (51.9)	65.18±33.71 (64.35)	75.43±33.35 (85.11)	0.647	33.84	0.55

*$P<0.05$, **$P<0.01$

HRV = heart rate variability; SD = Standard Deviation,

RR = RR-interval; ms = millisecond; HR = heart rate; bpm = beats per minute; SDNN = standard deviation of RR-interval; RMSSD = root mean square of successive differences in RR-intervals; pNN50 = percentage of intervals differing by >50 ms from preceding interval; SD1 = standard deviation of short term variability; SD2 = standard deviation of the long-term variability; HF = high-frequency components; LF = low-frequency components; LF/HF = autonomic balance; LF nu = low frequency power normalized units; HF nu = high frequency power normalized units;

CV = coefficient of variation; ICC = intraclass correlation coefficient.

5. Discussion

Overall, the present study, which investigated the effect of different CF for the automatic correction of artefacts in a tachogram, indicated that the HRV values obtained with the various CF produced different measures of statistical location (mean and median). The No, Low, Medium and Strong CF resulted in similar patterns of HRV as measured over the 5 day period. However, the median HRV values between Strong and Very Strong were overall significantly different from the rest. Thus, the Strong CF, which erased more artefacts and background noise without dampening the variability in the RR-interval signal, was selected as the most appropriate CF for this study.

The study also investigated the repeatability and reliability of HRV indicators in ponies in a restricted and unrestricted environment using the Strong CF. Heart rate measures showed poor repeatability, but the reliability was good in both environments. The HRV indicators showed good repeatability, but the reliability for the indicators were generally poor to good for the pasture and stocks data. The measurements obtained from the stocks were less reliable than those from the pasture. Using normalized low-frequency and normalized high-frequency components improved the repeatability and reliability of LF and HF.

5.1. The influence of different correction factors on heart rate measures and HRV indicator values

Errors or artefacts in heart rate variability data may occur due to factors relating to the normal physiology of a horse (pronounced T-wave, muscle contractions or movement of the horse [10], pathological conditions leading to disruption of electrical activity in the heart [49] and technical challenges associated with heart monitoring equipment [10, 31, 49]). Errors associated with the heart rate monitor are minimized by ensuring good electrode-skin contact, by visual exclusion of artefacts and correcting data mathematically with software packages.

Overall the results indicated that the HRV indicator values obtained with different CF produced different median values. Inconsistencies within the significant differences between MeanRR and MeanHR for the CF comparisons may be due to the reciprocal nature of these variables, that does not

allow them to be directly substituted for each other [50]. Results also showed that the use of No, Low and Medium CF during HRV quantification produced similar HRV indicator values.

The Kendall's Coefficient of Concordance indicated that the magnitude of the difference between the days were weak or very weak. This is consistent with the non-significant results obtained from the Friedman test, and therefore also the evidence of repeatability. The Strong vs. Very Strong CF indicated significant differences for most HRV indicators, as well as Strong and Very Strong vs. the rest of the factors (No, Low and Medium CF). It is very important to note that the graphical representation indicated that the Strong CF still followed the same pattern as the lower CF. The Very Strong CF on the other hand would have removed all day to day variability expected from an RR-interval series. Thus, in this study data set the Strong CF was able to strike a balance between the necessary removal of artefacts and background noise without removal of the variability of the RR-intervals.

Garza et al [37] compared Strong and Very Strong CF and excluded data obtained via the Very Strong CF due to "significant difference in output" between them, although the methodology followed or results were not described. Similar to the findings reported by Garza et al [37] the Very Strong CF was not appropriate to eliminate artefacts in the present study. It is thus important for within study repeatability to keep the CF consistent and to report the specific CF used, for study comparisons.

5.2. Repeatability of heart rate measures and HRV indicators

Repeatability studies must exclude bias between measurements, thus the within-subject standard deviation must agree on at least two measurements of the same subject [40]. One-way ANOVA performed on repeated measurements, and its non-parametric equivalent the Friedman test, can estimate the within-subject standard deviation.

Significant differences in heart rate measurements were only found in MeanRR and MeanHR on one day (Day 2) in both study environments. These differences could be explained by random changes in the "normal" environment on that day (i.e. a worker present during the pasture phase and a generator operating during the stocks phase). The overall results would thus suggest that there is a

good correlation (repeatability) between the various indicators of HRV as well as MeanRR and MeanHR.

5.3.   Reliability of heart rate measures and HRV indicators

Reliability indicates if the differences in the measurements are due to measurement error or due to normal variation [36]. In human sports medicine acceptable assessment of reliability of HRV includes ICC, CV and Limits of Agreement [29].

CV and ICC are used to evaluate the reliability of multiple repeated tests on an individual [35]. According to convention, the lower the value of CV, the more consistent the indicator performed over the monitoring period [39]. The CV of both pasture data and stocks data indicated low to good consistency with regards to the HRV indicators [29]. The CV obtained for LF, HF and LF/HF were markedly less consistent than the other indicators (stocks and pasture), whereas normalized LF and normalized HF were more consistent. Thus, the reliability for most of the indicators of HRV was poor to good based on the CV.

The interpretation of ICC, i.e. the translation of the values of ICC into categories indicating the degree of consistency, has not been proven [39]. In general, the nearer to 1 (one) the ICC value is, the better the relative reliability of the measurements and the nearer to 0 (zero) the poorer the relative reliability [51]. The ICC of the HRV indicators (pasture and stocks) appeared to be good, and in some instances poorly, reliable. Generally the CV values increased in data obtained from the stocks and the ICC values decreased compared to the data obtained from the pasture. Thus, the CV and ICC suggest that the HRV data obtained from the stocks setting were less reliable than the data obtained from the pasture setting. Using a stock to control the movement of the horse could therefore be considered as an intervention, with resultant effects on the HRV.

The data reported followed some of the trends of HRV results from humans, namely: using normalized LF and HF, rather than LF and HF improved the CV and ICC [29] and HRV is moderately to poorly reliable [25, 29] and in some instances good. Poor reliability in human HRV research was either due to an intervention or partly due to a poor experimental design [29]. Furthermore, in the

present study poor reliability of some of the HRV indicators was likely as horses cannot be controlled under experimental conditions to the same degree as humans. One of the few studies available on the reliability of HRV in horses was performed by comparing horses restricted to a stall and restricted in a stock [34]. Only a limited number of indicators were evaluated (MeanRR, SDRR, RMSSD, SD1 and SD2). Nonetheless, the results indicated that less restriction of movement in the stall appeared to be more reliable than restricted movement in the stocks. The results indicate that more "restrictive movement option" appeared to be less reliable, thus free movement seems to favor more reliable baseline HRV measurements. Unfortunately, unrestricted free walking may confound the interpretation of HRV measurements due to artefacts [34].

6. Conclusion

Graphical and statistical comparison showed that the specific CF used have an influence on the HRV indicator values. In this study the Strong CF was the most appropriate CF to use, due to the fact that the Very Strong CF resulted in very low day to day variability in the RR-interval signal, and after removal of artefacts and background noise the Strong CF still followed the same pattern as the No, Low and Medium CF. This information may be useful during application of HRV quantification in similar studies. Balancing the advantages and disadvantages of choosing different CF for RR-interval series during HRV quantification and reporting the CF applied is vital to increase comparability between studies.

The time domain, frequency domain and geometric indicators did not differ significantly over the five day period during free movement on pasture or during restrictive movement in examination stocks, suggesting that HRV indicators have good repeatability. The measurements obtained in the examination stocks were less reliable than those obtained on pasture and the frequency domain indicators could be improved by using normalized LF and HF. Overall, the reliability of HRV indicators were similar to findings in human studies. Finally, the choice of the CF used in HRV studies should be carefully considered in each setting it is applied to and the choice should be clearly defined and defendable so as to standardize the procedures used in HRV data analysis.

7. Acknowledgments

8. References

[1]   Pell SM, McGreevy PD. A study of cortisol and beta-endorphin levels in stereotypic and normal Thoroughbreds. Appl Anim Behav Sci 1999;64:81-90.

[2]   Dzikiti T, Hellebrekers L, Dijk P. Effects of intravenous lidocaine on isoflurane concentration, physiological parameters, metabolic parameters and stress-related hormones in horses undergoing surgery. J Vet Med Ser A 2003;50:190-5.

[3]   Verbeek E, Oliver MH, Waas JR, McLeay LM, Blache D, Matthews LR. Reduced cortisol and metabolic responses of thin ewes to an acute cold challenge in mid-pregnancy: implications for animal physiology and welfare. PLoS ONE 2012;7:e37315.

[4]   Moberg G. Biological response to stress: implications for animal welfare. The biology of animal stress: Basic principles and implications for animal welfare CABI Publishing, New York 2000:1-21.

[5]   Schmidt A, Biau S, Möstl E, Becker-Birck M, Morillon B, Aurich J, et al. Changes in cortisol release and heart rate variability in sport horses during long-distance road transport. Domest Anim Endocrinol 2010;38:179-89.

[6]   Schmidt A, Möstl E, Wehnert C, Aurich J, Müller J, Aurich C. Cortisol release and heart rate variability in horses during road transport. Horm Behav 2010;57:209-15.

[7]   Christensen JW, Malmkvist J, Nielsen BL, Keeling L. Effects of a calm companion on fear reactions in naive test horses. Equine Vet J 2008;40:46-50.

[8]   Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Circulation 1996;93:1043-65.

[9]   Grant CC, Clark JR, van Rensburg DCJ, Viljoen M. Relationship between exercise capacity and heart rate variability: Supine and in response to an orthostatic stressor. Auton Neurosci 2009;151:186-8.

[10]   von Borell E, Langbein J, Despres G, Hansen S, Leterrier C, Marchant-Forde J, et al. Heart rate variability as a measure of autonomic regulation of cardiac activity for assessing stress and welfare in farm animals - A review. Physiol Behav 2007;92:293-316.

[11]   Visser EK, van Reenen CG, van der Werf JTN, Schilder MBH, Knaap JH, Barneveld A, et al. Heart rate and heart rate variability during a novel object test and a handling test in young horses. Physiol Behav 2002;76:289-96.

[12]   Erber R, Wulf M, Aurich J, Rose-Meierhöfer S, Hoffmann G, von Lewinski M, et al. Stress response of three-year-old horse mares to changes in husbandry system during initial equestrian training. J Equine Vet Sci 2013;33:1088-94.

[13]   Stachurska A, Janczarek I, Wilk I, Kędzierski W. Does Music Influence Emotional State in Race Horses? J Equine Vet Sci 2015;35:650-6.

[14]   Werhahn H, Hessel EF, Van den Weghe HF. Competition Horses Housed in Single Stalls (II): Effects of Free Exercise on the Behavior in the Stable, the Behavior during Training, and the Degree of Stress. J Equine Vet Sci 2012;32:22-31.

[15]   Mésangeau D, Laude D, Elghozi J-L. Early detection of cardiovascular autonomic neuropathy in diabetic pigs using blood pressure and heart rate variability. Cardiovasc Res 2000;45:889-99.

[16] Voss L, Bolton D, Galland B, Taylor B. Endotoxin effects on markers of autonomic nervous system function in the piglet: implications for SIDS. Neonatology 2004;86:39-47.

[17] Désiré L, Veissier I, Després G, Boissy A. On the way to assess emotions in animals: do lambs (Ovis aries) evaluate an event through its suddenness, novelty, or unpredictability? J Comp Psychol 2004;118:363-74.

[18] Langbein J, Nürnberg G, Manteuffel G. Visual discrimination learning in dwarf goats and associated changes in heart rate and heart rate variability. Physiol Behav 2004;82:601-9.

[19] Coulon M, Hild S, Schroeer A, Janczak A, Zanella A. Gentle vs. aversive handling of pregnant ewes: II. Physiology and behavior of the lambs. Physiol Behav 2011;103:575-84.

[20] Mohr E, Langbein J, Nürnberg G. Heart rate variability: a noninvasive approach to measure stress in calves and cows. Physiol Behav 2002;75:251-9.

[21] Stewart M, Stafford K, Dowling S, Schaefer A, Webster J. Eye temperature and heart rate variability of calves disbudded with or without local anaesthetic. Physiol Behav 2008;93:789-97.

[22] Korte SM, Ruesink W, Blokhuis HJ. Heart rate variability during manual restraint in chicks from high-and low-feather pecking lines of laying hens. Physiol Behav 1998;65:649-52.

[23] Hull Jr SS, Evans AR, Vanoli E, Adamson PB, Stramba-Badiale M, Albert DE, et al. Heart rate variability before and after myocardial infarction in conscious dogs at high and low risk of sudden death. J Am Coll Cardiol 1990;16:978-85.

[24] Calvert CA, Jacobs GJ. Heart rate variability in Doberman Pinschers with and without echocardiographic evidence of dilated cardiomyopathy. Am J Vet Res 2000;61:506-11.

[25] Lord S, Senior R, Das M, Whittam A, Murray A, McComb J. Low-frequency heart rate variability: reproducibility in cardiac transplant recipients and normal subjects. Clin Sci 2001;100:43-6.

[26] Sookan T, McKune AJ. Heart rate variability in physically active individuals: reliability and gender characteristics. Cardiovasc J Afr 2012;23:67.

[27] Cipryan L, Litschmannova M. Intra-day and inter-day reliability of heart rate variability measurement. J Sports Sci 2013;31:150-8.

[28] Tannus LR, Sperandei S, Montenegro Júnior RM, Carvalho VR, Pedrosa HC, Félix MT, et al. Reproducibility of methods used for the assessment of autonomous nervous system's function. Auton Neurosci 2013;177:275-9.

[29] Sandercock GRH, Bromley PD, Brodie DA. The reliability of short-term measurements of heart rate variability. Int J Cardiol 2005;103:238-47.

[30] Stucke D, Große Ruse M, Lebelt D. Measuring heart rate variability in horses to investigate the autonomic nervous system activity – Pros and cons of different methods. Appl Anim Behav Sci 2015;166:1-10.

[31] Parker M, Goodwin D, Eager RA, Redhead ES, Marlin DJ. Comparison of Polar® heart rate interval data with simultaneously recorded ECG signals in horses. J Exerc Physiol 2009;6:137-42.

[32] Kato T, Ohmura H, Hiraga A, Wada S, Kuwahara M, Tsubone H. Changes in heart rate variability in horses during immersion in warm springwater. Am J Vet Res 2003;64:1482-5.

[33] Kinnunen S, Laukkanen R, Haldi J, Hanninen O, Atalay M. Heart rate variability in trotters during different training periods. Equine Vet J 2006;38:214-7.

[34] Vitale V, Balocchi R, Varanini M, Sgorbini M, Macerata A, Sighieri C, et al. The effects of restriction of movement on the reliability of heart rate variability measurements in the horse (Equus caballus). J Vet Behav 2013;8:400-3.

[35] Smiet E, Van Dierendonck MC, Sleutjens J, Menheere PPCA, van Breda E, de Boer D, et al. Effect of different head and neck positions on behaviour, heart rate variability and cortisol levels in lunged Royal Dutch Sport horses. The Veterinary Journal 2014;202:26-32.

[36] Sundra T, Harrison J, Lester G, Raidal S, Phillips J. The influence of spasmolytic agents on heart rate variability and gastrointestinal motility in normal horses. Res Vet Sci 2012;93:1426-33.

[37] Garza JL, Cavallari JM, Eijckelhof BH, Huysmans MA, Thamsuwan O, Johnson PW, et al. Office workers with high effort–reward imbalance and overcommitment have greater decreases in heart rate variability over a 2-h working period. Int Arch Occup Environ Health 2014:1-11.

[38] Quan H, Shih WJ. Assessing reproducibility by the within-subject coefficient of variation with random effects models. Biometrics 1996:1195-203.

[39] Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med 1998;26:217-38.

[40] Bartlett J, Frost C. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. Ultrasound Obstet Gynecol 2008;31:466-75.

[41] Hopkins WG. Measures of reliability in sports medicine and science. Sports Med 2000;30:1-15.

[42] Ohmura H, Hiraga A, Aida H, Kuwahara M, Tsubone H. Effects of repeated atropine injection on heart rate variability in Thoroughbred horses. J Vet Med Sci 2001;63:1359-60.

[43] Cottin F, Médigue C, Lopes P, Petit E, Papelier Y, Billat V. Effect of exercise intensity and repetition on heart rate variability during training in elite trotting horse. Int J Sports Med 2005;26:859-67.

[44] Hada T, Mukai HOK, Eto D, Takahashi T, Hiraga A. Utilisation of the time constant calculated from heart rate recovery after exercise for evaluation of autonomic activity in horses. Equine Vet J 2006;38:141-5.

[45] Gehrke EK, Baldwin A, Schiltz PM. Heart rate variability in horses engaged in equine-assisted activities. J Equine Vet Sci 2011;31:78-84.

[46] Mourot L, Bouhaddi M, Perrey S, Cappelle S, Henriet M-T, Wolf J-P, et al. Decrease in heart rate variability with overtraining: assessment by the Poincaré plot analysis. Clin Physiol Funct Imaging 2004;24:10-8.

[47] Tarvainen MP, Niskanen JP. Kubios HRV Version 2.1 Users Guide: Department of Applied Physics, University of Eastern Finland, Kuopio, Finland; 2012.

[48] Ille N, Erber R, Aurich C, Aurich J. Comparison of heart rate and heart rate variability obtained by heart rate monitors and simultaneously recorded electrocardiogram signals in non-exercising horses. J Vet Behav 2014;9:341-6.

[49] Marchant-Forde R, Marlin D, Marchant-Forde J. Validation of a cardiac monitor for measuring heart rate variability in adult female pigs: accuracy, artefacts and editing. Physiol Behav 2004;80:449-58.

[50] Goldberger JJ, Johnson NP, Subacius H, Ng J, Greenland P. Comparison of the physiologic and prognostic implications of the heart rate versus the RR interval. Heart Rhythm 2014;11:1925-33.

[51] Pinna G, Maestri R, Torunski A, Danilowicz-Szymanowicz L, Szwoch M, La Rovere M, et al. Heart rate variability measures: a fresh look at reliability. Clin Sci 2007;113:131-40.