Original Article

Intra-session stability of short-term heart rate variability measurement: gender and total spectral power Influence

LUKAS CIPRYAN [1], MARTINA LITSCHMANNOVA [2]

[1] *Ostrava University, Human Motion Diagnostic Centre*

[2] *VSB-Technical University of Ostrava, Faculty of Electrical Engineering and Computer Science, Department of Applied Mathematics*

ABSTRACT

Cipryan, L., & Litschmannova, M. (2014). Intra-session Stability of Short-term Heart Rate Variability Measurement: Gender and Total Spectral Power Influence. *J. Hum. Sport Exerc., 9*(1), pp.68-80. Heart rate variability (HRV) has been increasingly analysed under numerous research settings. HRV measurement reliability is, however, still an unresolved issue. The main purpose is to carry out an intra-session stability evaluation of HRV parameters from short-term recordings by means of orthoclinostatic stimulation in a study group which is stratified by gender or Total Power (PT) magnitude. The goal is to make as homogeneous a study group as possible and investigate whether the reproducibility level could be influenced by these factors. The study group consisted of 103 participants (age 22.3 ± 1.2). Standard HRV indexes were computed: PT (total spectral power), PHF (high frequency spectral power), PLF (low frequency spectral power) and LF/HF. Absolute reliability is assessed by the standard error of measurement and 95% limits of agreement; the relative reliability is assessed by the intraclass correlation coefficient. The markedly different standard error of measurement (SEM) between the Male and Female groups was not observed for any HRV parameters. The intraclass correlation coefficient (ICC) values ranged from 0.67 to 0.95 for males and from 0.69 to 0.97 for females. According to the SEM and ICC, there is no difference between the groups of High PT and Low PT. There are not any significant differences in absolute or relative reliability between the more homogeneous study groups and we have therefore concluded that HRV measurement reliability is not influenced by gender or HRV magnitude. **Key words**: HEART RATE VARIABILITY, RELIABILITY, AUTONOMIC NERVOUS SYSTEM, TEST-RETEST, ORTHOCLINOSTATIC

Corresponding author. Ostravska univerzita, CDLP. Varenska 40a. 702 00 Ostrava. Czech Republic.
E-mail: lukas.cipryan@osu.cz
Submitted for publication July 2013
Accepted for publication January 2014
JOURNAL OF HUMAN SPORT & EXERCISE ISSN 1988-5202

doi:10.4100/jhse.2014.91.08

INTRODUCTION

Heart rate variability (HRV) is a non-invasive diagnostic tool for cardiac autonomic regulation evaluation. This indicator has been increasingly analysed in a wide range of research and clinical settings, including sport sciences as well. HRV can be influenced by various external or internal factors and there are also strong connections with certain diseases such as Metabolic syndrome (Koskinen et al., 2009; Stein et al., 2007) and additional cardiovascular diseases (Brook & Julius, 2000; Dekker et al., 2000; Malpas, 2002). Age (Kobayashi, 2007; Migliaro et al., 2001), physical condition (Carter et al., 2003), mental stress (Hjortskov et al., 2004), sport training (Billman, 2002; Sandercock et al., 2005) and heart rate (Nieminen et al., 2007) rank among the most significant factors with a direct effect on HRV.

Gender is also considered an additional meaningful factor. It influences HRV less in comparison with age. Both factors, however, are important determinants of HRV in healthy individuals (Jensen-Urstad et al., 1997). Britton et al. (2007) present higher values of parasympathetic activity variables and lower values of sympathetic activity variables in women than in men of all age categories. In addition, natural HRV decreasing with ageing is more pronounced in men. Additional studies have suggested that gender differences diminish at the age of 40 (Ramaekers et al., 1998), at 60 years of age (Kuo et al., 1999) or earlier at 30 (Umetani et al., 1998).

Since the original publication of the HRV Task Force standards in 1996, a variety of new methods and procedures has been proposed to quantify HRV. As regards spectral analysis, HRV measurement validity and spectral component interpretation are relatively stable. The high frequency spectral component (HF), which is under respiratory influence, is considered to have been exclusively modulated by the efferent cardiac vagal activity (Martinmäki et al., 2006; Task Force 1996). Physiological correlates and the biological relevance of the low frequency (LF) spectral component are more controversial. The LF component is believed by some to be a marker of sympathetic modulation and by others as a parameter that includes both sympathetic and vagal influences (De Meersman & Stein, 2007). The physiological interpretation of the very low or ultra low spectral components is still largely unknown.

The question of HRV measurement reliability, or intra-session stability in our case, is, however, continually accompanied by numerous doubts. Various methodological approaches to HRV measurement and analysis (short-term vs. long-term; the type of cardiovascular stimulation; time vs. the frequency domain evaluation methods, etc.) are the most probable reasons. This study focuses on two questions, which have not been discussed and published yet, as far as we are aware. First of all, it is assumed that there are significant differences in cardiac autonomic regulation between males and females. We are interested in investigating if this fact might be reflected in the back-to-back reproducibility level as well. The second research purpose is similar. We would like to clarify if there are any differences in HRV measurement stability between individuals with high and low HRV. More homogeneous study groups have been, therefore, created in order to eliminate the possible impact of heterogeneity of the study group on the intra-session stability level. We build upon our previous work (Cipryan & Litschmannova, 2013), in which we demonstrated low absolute reliability regardless of gender or HRV magnitude. The goal of our research efforts is to continue to seek out a more reliable means of HRV examination and demonstrate its influencing factors.

MATERIAL AND METHODS

Participants

We studied 103 volunteers (age 22.3 ± 1.2; 54 males, 49 females). They were without any acute health disorders at the time of the HRV measurement, evidencing normal levels of blood pressure and ECG patterns. They were non-obese, took no medication or other dietary supplements and were non-smokers. All the participants were routinely involved in various sport activities at the recreational level. None had been involved in any regular sport training for at least 6 months. The study conformed with the recommendations of the Declaration of Helsinki. Written consent to participate in this experiment, which was approved by the local ethics committee, was obtained from each individual.

Research design

Orthoclinostatic stimulation was used for the short-term HRV examination, which we repeated without any interruption of the measurement (measurement 1 – measurement 2; M1vsM2). The HRV measurement was performed at approximately the same time of day for all the individuals (7:30 – 10:00 am.) and carried out under laboratory conditions (a dimly lit, quiet, climate controlled room).

The participants were informed about the research conditions and were encouraged to adhere to all standard requirements, i.e. avoid any intensive physical activity for 48 hours, eating, caffeine and alcohol drinks before the HRV measurement. Verbal confirmation in connection with the observance of these conditions was required immediately prior to the trials. No participants broke these rules.

The established orthoclinostatic stimulation consisted of three intervals: supine – standing – supine. The first interval (supine) serves the function of the preceding relaxation period and is not included in the HRV analysis. This was the reason for removing the first interval from the retest measurement, in order not to repeat two successive supine intervals. The test-retest procedure was consequently performed in five intervals (i.e. supine – standing – supine – standing – supine). Each interval lasted 5 minutes, which is the preferred and recommended ECG recording duration for a short-term HRV analysis (Task Force, 1996). The position change was active. When the interval was completed, a short acoustic signal (one beep) was emitted by the PC (part of the software for HRV measurement) and the participant promptly stood up (remaining unsupported for the entire interval) or lay down on a laboratory bed. The procedure was always supervised by a trained and experienced researcher. Breath frequency was not conducted.

HRV analysis

ECG was sampled at 1000 Hz with VarCor PF7 (Dimea Group Ltd, Olomouc, Czech Republic; Salinger and Gwozdziewicz, 2008). This diagnostic system enables a routine short-term HRV evaluation with respect to Task Force (1996) findings and recommendations. The accuracy of the measurements was 1 ms. The short-term recording of the RR interval (the duration between two consecutive R peaks) was visually validated prior to the frequency analysis. The presence of ectopy, missing data or noise was eliminated. Due to non-stationarity, there was 1 minute between each interval (the phase of the changing position), which was not included in the spectral analysis of HRV. The power spectrum was obtained by transforming the time data (the duration of the RR intervals) into frequency values. The spectral density was calculated by means of the Fast Fourier Transform with partially adapted CGSA procedures (Yamamoto et al., 1991). This algorithm assures optimum suppression of the non-harmonics and noise components of the analysed signal. Three main spectral components were distinguished: a very low frequency (VLF) $\leq$ 0.04 Hz, a low frequency (LF) 0.04-0.15 Hz and a high frequency (HF) 0.15-0.40 Hz (Task Force, 1996). The spectral variables Total spectral power (PT), LF power (PLF), HF power (PHF), ratio LF/HF and LF power in normalized units (LFnu) were employed in this research. To avoid redundancy and keep the study comprehensible, the analysis was based on these most frequently used parameters exclusively.

Statistical analysis

In order to carry out the aims of this research, the study group was divided into these subgroups:

1. Males (n = 54, age 22.6 ± 1.3, min - max: 20.6 - 25.8) vs. Females (n = 49, age 22.1 ± 1.1, min - max: 20.5 - 25.3),
2. High PT (n = 40, age 22.4 ± 1.0, min - max: 20.6 - 24.7) vs. Low PT (n = 40, age 22.2 ± 1.2, min - max: 20.8 - 25.1). All the results of the HRV measurements were lined up according to the absolute value of the Total spectral power (PT) in the third supine position (5th interval). The two subgroups with the highest and lowest PT were consequently created.

Prior to measurement 1 (M1) - measurement 2 (M2) paired data were compared, the collected data were checked to detect outliers and to verify sampling distribution and the possible presence of heteroscedasticity. Normal distribution and homogeneity of variance were verified by the Lilliefors test and plotting the Bland-Altman plot and scatter plot, respectively. All the variables (apart from the Mean RR) revealed non-normality and heteroscedasticity, therefore the data were log-transformed using the natural logarithm (ln). The log-transformation was effective in decreasing heteroscedasticity (see Figure 1). An observed (or log-transformed) value was detected as an outlier if it was less/greater than the lower/upper quartile ± 1.5 times the interquartile range. The outliers were removed and not included in the statistical analysis. Standard statistical methods were used to calculate the basic descriptive statistics for all the variables in the following step. The zero mean of the difference between the two measurements (two-sided paired Student's t-test at the 0.05 significance level) was also tested.

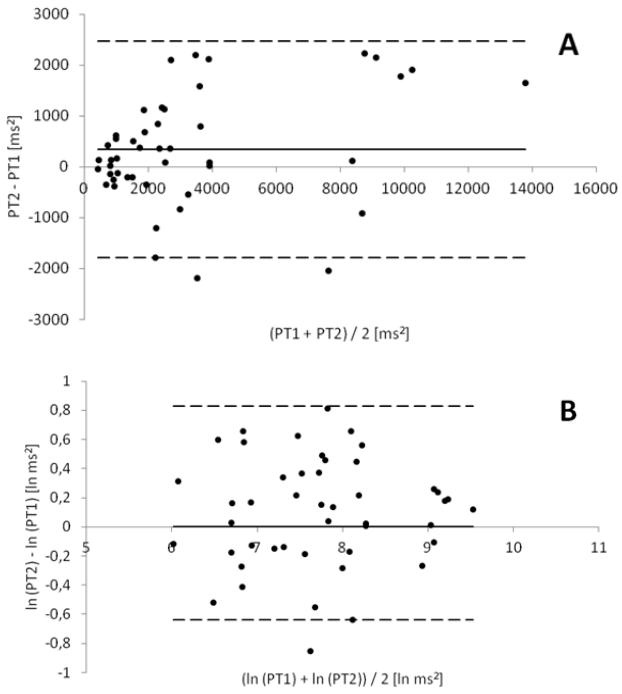

Figure 1. *Bland-Altman plots for the chosen HRV parameter (PT, females) before (A) and after (B) logarithmic transformation. A systematic change between the two measurements (non-symmetrical dispersion around the zero line) and heteroscedasticity (the magnitude of scattering around the zero line steadily increases) were eliminated and normal data distribution and homoscedasticity were produced.*

According to Pinna et al. (2007), the absolute and relative expression of the test-retest stability of the HRV measurement was assessed. A key statistical indicator of absolute reliability is the standard error of

measurement (SEM), calculated in this study according to Strauss et al. (2006). We also computed the Effect size (ES), which represents the ratio of the mean difference over the pooled standard deviation and was used to estimate the magnitude of the analysed differences. The Bland and Altman plots, as well as 95% limits of agreement (95% LoA; Bland and Altman, 1986), were also used for the assessing agreement between the two measurements. 95% LoA was estimated by a mean difference of 1.96 standard deviation of the differences. When the Bland Altman plot and 95% LoA for the log-transformed data were computed, we transformed these limits of agreement back to the original scale by taking anti-logs. This yields an interval for the ratio between the two measurements. 95% LoA indicates that a significant change has occurred, if the observed difference (retest minus test) or ratio (retest / test) lies outside this interval (with 95% probability) (Atkinson and Nevill, 1998; Bland and Altman, 1999).

The relative intra-session stability was evaluated by an intraclass correlation coefficient (ICC). ICC and confidence intervals (with 95% probability) for ICC were calculated according to Hopkins (2009).

Finally, a sample size needed to detect a relevant change in the mean of the HRV parameters was estimated. A change in the mean of $\geq$ 30 % of between-individual standard deviation is, conventionally, considered "relevant".

The data was statistically analysed with Microsoft® Office Excel 2007 and IBM SPSS Statistics 20.

RESULTS

The statistical analysis did not show any difference between HRV parameters stability in standing and supine. Therefore, the results are presented only for the supine position in order to keep the study as simple and understandable as possible.

Gender influence

All the HRV variables (apart from the Mean RR) indicate non-normality and heteroscedasticity and were consequently logarithmically transformed. According to the paired Student's t-test for the HRV parameters in the supine position, LF/HF and LFnu (both in males and females), and also PT and PLF in females manifest no significant ($P > 0.05$) test-retest mean difference (Table 1). The diverse magnitude of the mean difference described by ES (Table 2) is apparent between males and females in PLF, LF/HF and LFnu.

Intra-session stability indexes are presented in Table 2. When we consider the absolute intra-session stability expressed by SEM, observable differences between males and females are not found. When we examine the limits of agreement (LoA), which is the only supporting index not revealing the magnitude of reliability, there is also no significant difference between males and females. The narrowest LoA interval is detected for PT and PHF in both groups of males and females. The worst case is found in the PLF (males), in which the second measurement can be as large / small as 7.14 / 0.29 times the first measurement due to pure random variation.

The ICC values ranged from 0.67 to 0.95 for males and from 0.72 to 0.97 for females. This indicates that random error accounted for approximately 30 % maximum of the total measurement variability. This is the same for both groups. The highest ICC was computed for Mean RR, PT and PHF (Table 2).

The estimated sample size needed to detect a significant test-retest change in the mean of the HRV parameters ranged from 4 to 32 and was extremely similar between males and females.

Table 1. *Descriptive results for HRV parameters – SUPINE position*
Values are expressed as mean (standard deviation). Skewed variables are also reported after log transformation.

	n	M1	M2	Difference (M2-M1)	P value (M1vsM2)
MALES					
Mean RR (ms)	52	1077 (151)	1106 (154)	29 (37)	0.00
PT ln (ln ms^2)	49	7.90 (0.99)	8.08 (0.93)	0.17 (0.47)	0.01
PLF ln (ln ms^2)	45	5.94 (1.07)	6.30 (1.09)	0.36 (0.82)	0.01
PHF ln (ln ms^2)	47	7.31 (1.11)	7.50 (1.05)	0.18 (0.41)	0.00
LF/HF ln	45	-1.55 (1.14)	-1.36 (1.05)	0.19 (0.76)	0.10
LFnu ln (%)	52	2.93 (0.93)	3.09 (0.86)	0.16 (0.63)	0.08
FEMALES					
Mean RR (ms)	48	952 (156)	974 (155)	22 (28)	0.00
PT ln (ln ms^2)	46	7.66 (0.90)	7.76 (0.91)	0.10 (0.37)	0.09
PLF ln (ln ms^2)	39	5.51 (0.87)	5.54 (0.91)	0.03 (0.68)	0.80
PHF ln (ln ms^2)	45	7.15 (1.15)	7.32 (1.15)	0.17 (0.40)	0.01
LF/HF ln	43	-1.75 (1.23)	-1.81 (1.09)	-0.07 (0.80)	0.59
LFnu ln (%)	47	2.74 (1.08)	2.68 (0.96)	-0.06 (0.69)	0.58

Legend: n – number of participants (without outliers); M1 / M2 – measurement 1 / 2; P value – significance level of paired Student´s t-test.

Table 2. *Intra-session stability of HRV parameters – SUPINE position*

	SEM	95% LoA	ES	ICC (95% CI)	Change in the mean to be detected (required sample size)
Males					
Mean RR	25.90 ms	(29.00 ± 72.52) ms	0.80	0.95 (0.92-0.97)	64.20 ms (n=5)
PT ln	0.33 ln ms^2	(1.19 */÷ 2.51)	0.37	0.87 (0.78-0.92)	0.40 ln ms^2 (n=14)
PLF ln	0.58 ln ms^2	(1.43 */÷ 4.99)	0.44	0.67 (0.48-0.81)	0.42 ln ms^2 (n=32)
PHF ln	0.29 ln ms^2	(1.20 */÷ 2.23)	0.44	0.92 (0.85-0.95)	0.45 ln ms^2 (n=9)
LF/HF ln	0.54 ln	(1.21 */÷ 4.44)	0.25	0.75 (0.59-0.86)	0.44 ln (n=26)
LFnu ln	0.44 ln	(1.17 */÷ 3.44)	0.25	0.75 (0.60-0.85)	0.36 ln (n=27)
Females					
Mean RR	19.90 ms	(22.00 ± 54.88) ms	0.78	0.97 (0.95-0.99)	65.60 ms (n=4)
PT ln	0.26 ln ms^2	(1.11 */÷ 2.07)	0.26	0.91 (0.84-0.95)	0.38 ln ms^2 (n=10)
PLF ln	0.48 ln ms^2	(1.03 */÷ 3.79)	0.04	0.72 (0.52-0.84)	0.35 ln ms^2 (n=32)
PHF ln	0.29 ln ms^2	(1.19 */÷ 2.19)	0.42	0.93 (0.87-0.96)	0.48 ln ms^2 (n=8)
LF/HF ln	0.56 ln	(0.93 */÷ 4.80)	0.08	0.77 (0.61-0.87)	0.46 ln (n=26)
LFnu ln	0.49 ln	(0.94 */÷ 3.87)	0.08	0.77 (0.63-0.87)	0.41 ln (n=25)

Legend: SEM – standard error of measurement; LoA – limits of agreement for the difference (only Mean RR) or ratio (all log transformed data) between the two measurements; ES – effect size; ICC – intraclass correlation coefficient; CI – confidence interval.

HRV magnitude influence

Similarly to the previous part, the HRV variables (apart from the Mean RR) were logarithmically transformed prior to statistical analysis. All HRV variables (apart from the Mean RR) in the supine position manifest no significant ($P > 0.05$) test-retest mean difference in the Low PT group. There is, however, statistically significant zero mean difference ($P \leq 0.05$) for all HRV variables in the High PT group (Table 3). This is also supported by the lower ES in the Low PT group (Table 4).

The most important statistical expression of absolute reliability is, however, the SEM. According to the SEM, there is no difference between the groups of High PT and Low PT. The LoA intervals are similarly wide for PT and PHF in both groups. The marked difference of the LoA interval extent between High PT and Low PT groups is found in the PLF and its related variables LF/HF or LFnu in the supine position. The worst case is found in the PLF parameter (High PT), in which the second measurement can be as large / small as 8.38 / 0.34 times the first measurement due to pure random variation (Table 4).

There is no obvious difference in the ICC between the groups. The ICC values ranged from 0.56 to 0.93 for High PT groups and from 0.62 to 0.97 for the Low PT group. This means that random error accounted for 44 % to 7 % and for 38 % to 3 %, respectively, of the total measurement variability. The highest ICC, always exceeding the value 0.90, was computed in all cases for Mean RR. The most obvious difference in ICC is found for PLF (0.56 vs. 0.72). The ICC differences are negligible in other HRV variables (Table 4).
The estimated sample size needed to detect a significant test-retest change in the mean of the HRV parameters ranged from 6 to 39 (High PT) and from 4 to 43 (Low PT).

Table 3. *Descriptive results for HRV parameters – SUPINE position*
Values are expressed as mean (standard deviation). Skewed variables are also reported after log transformation.

	n	M1	M2	Difference (M2-M1)	P value (M1vsM2)
HIGH PT					
Mean RR (ms)	38	1114 (115)	1143 (122)	30 (34)	0.00
PT ln (ln ms^2)	39	8.76 (0.66)	8.97 (0.47)	0.21 (0.41)	0.00
PLF ln (ln ms^2)	34	6.54 (1.08)	7.06 (0.88)	0.52 (0.82)	0.00
PHF ln (ln ms^2)	37	8.32 (0.79)	8.53 (0.59)	0.20 (0.38)	0.00
LF/HF ln	31	-1.98 (1.15)	-1.68 (1.00)	0.29 (0.79)	0.05
LFnu ln (%)	33	2.56 (1.07)	2.83 (0.93)	0.27 (0.67)	0.03
LOW PT					
Mean RR (ms)	40	923 (156)	951 (156)	24 (32)	0.00
PT ln (ln ms^2)	36	6.90 (0.53)	7.00 (0.51)	0.10 (0.45)	0.18
PLF ln (ln ms^2)	35	5.48 (0.75)	5.41 (0.86)	-0.07 (0.66)	0.52
PHF ln (ln ms^2)	36	6.23 (0.82)	6.36 (0.80)	0.12 (0.44)	0.10
LF/HF ln	34	-1.05 (1.11)	-1.22 (1.09)	-0.17 (0.76)	0.20
LFnu ln (%)	38	3.27 (0.87)	3.14 (0.87)	-0.13 (0.63)	0.21

Legend: n – number of participants (without outliers); M1 / M2 – measurement 1 / 2; P value – significance level of paired Student´s t-test.

Table 4. *Intra-session stability of HRV parameters – SUPINE position*

	SEM	95% LoA	ES	ICC (95% CI)	Change in the mean to be detected (required sample size)
HIGH PT					
Mean RR	24.20 ms	(30.00 ± 66.64) ms	0.87	0.93 (0.87-0.96)	49.80 ms (n=6)
PT ln	0.29 ln ms^2	(1.23 */÷ 2.23)	0.51	0.69 (0.49-0.83)	0.23 ln ms^2 (n=28)
PLF ln	0.58 ln ms^2	(1.68 */÷ 4.99)	0.64	0.56 (0.27-0.75)	0.38 ln ms^2 (n=39)
PHF ln	0.27 ln ms^2	(1.22 */÷ 2.11)	0.54	0.81 (0.67-0.90)	0.28 ln ms^2 (n=17)
LF/HF ln	0.56 ln	(1.34 */÷ 4.70)	0.37	0.71 (0.48-0.85)	0.43 ln (n=30)
LFnu ln	0.47 ln	(1.31 */÷ 3.72)	0.41	0.75 (0.56-0.87)	0.40 ln (n=24)
LOW PT					
Mean RR	22.70 ms	(24.00 ± 62.72) ms	0.73	0.97 (0.94-0.98)	65.90 ms (n=4)
PT ln	0.32 ln ms^2	(1.11 */÷ 2.42)	0.23	0.62 (0.36-0.78)	0.20 ln ms^2 (n=43)
PLF ln	0.48 ln ms^2	(0.93 */÷ 3.65)	0.04	0.72 (0.52-0.84)	0.35 ln ms^2 (n=32)
PHF ln	0.31 ln ms^2	(1.13 */÷ 2.37)	0.29	0.85 (0.72-0.92)	0.33 ln ms^2 (n=16)
LF/HF ln	0.54 ln	(0.84 */÷ 4.44)	0.23	0.75 (0.56-0.87)	0.44 ln (n=26)
LFnu ln	0.45 ln	(0.88 */÷ 3.44)	0.21	0.73 (0.54-0.85)	0.34 ln (n=29)

Legend: SEM – standard error of measurement; LoA – limits of agreement for the difference (only Mean RR) or ratio (all log transformed data) between the two measurements; ES – effect size; ICC – intraclass correlation coefficient; CI – confidence interval.

DISCUSSION

Intra-session stability estimation

The statistical procedure in this study is based on the recommendation for reliability investigation, even if we term our problem as intra-session stability. Retest reliability refers to the reproducibility of a measurement when it is repeated for a reasonable number of times on a reasonable number of individuals (Hopkins, 2000). The retest repetition has to be conducted, as much as possible, under identical conditions. The possible influence of each measurement on the consequent measurements has to be simultaneously eliminated. This is the reason why we cannot refer to the study subject as a genuine reliability study. The unchanging conditions were met although the possible influence of the first measurement on the second measurement (e.g. the change in the level of vigilance) could have appeared in the non-interrupted test-retest procedure. In contrast, we have previously demonstrated that there is not a reliability difference between intra-session and inter-day HRV measurements (Cipryan & Litschmannova, 2013).

Within-individual variation is the most important type of reliability measurement as it affects the precision of estimates of change in the variable of an experimental study. The standard deviation of the individual's value or the standard error of measurement is a statistic which captures the random variability of a single individual's values with repeated measuring (Hopkins, 2000). Additional supporting statistical indicators are also included in order to achieve as complete an HRV reliability picture as possible. The most important parameter SEM cannot be, however, replaced by these. The Bland-Altman 95% LoA, presented in Tables 2

and 4, identify the range of possible systematic change between the two measurements. Similar information is indicated by the paired Student's t-test and Effect Size.

The evaluation of the relative reliability is based on a correlation analysis. This dimensionless type of measurement presents how closely the values of one trial track the values of another as we move our attention from individual to individual (Hopkins, 2000). The sufficient relative reliability is usually considered for a correlation coefficient higher than 0.80 and a substantial reliability for ICC between 0.60 – 0.80. There is a need to be aware of the fact that the correlation analysis is extremely sensitive to the range of values in the sample. Large between-individual variability produces high values of correlation coefficients (Atkinson & Nevill, 1998).

Gender influence on intra-session stability

Although HRV depends on gender, the significant gender-related difference of HRV decreases with ageing (Bonnemeier et al., 2003). Antelmi et al. (2004) have observed that the sympathovagal parameters are higher in males, whereas HRV parameters representing efferent vagal modulation are higher in females. Likewise, Britton et al. (2007) demonstrate higher PLF values in males and higher PHF values in females (see also Hedelin et al., 2000). According to Huikuri et al. (1996), baroreflex responsiveness is attenuated in middle-aged women as compared with men, although the cardiac vagal modulation is augmented. Hormone replacement therapy (and also physical activity – see Davy et al., 1996) appears to have favourable effects on cardiovascular autonomic regulation in postmenopausal women (Kuch et al., 2001).

The normal cyclic variations in endogenous sex hormone levels during the menstrual cycle are not significantly associated with changes in cardiac autonomic control as measured by HRV. A significant correlation between peak estrogen levels and HRV measures at ovulation provides further support for the reported cardioprotective effects of estrogen in healthy females (Leicht et al., 2003). We did not monitor the possible impact of the menstrual cycle in this study, which can be considered as a study limit. We are convinced, however, that it may have more influence on the HRV results than the level of HRV measurement reliability.

It is apparent that HRV differences between males and females exist. This study aims, however, to conclude if HRV measurement reliability depends on gender, since this is one of the important factors influencing HRV. The unsatisfactory absolute reliability, expressed by SEM, has been already demonstrated regardless of the attempt to make the study group more homogeneous (Cipryan & Litschmannova, 2013). The study group heterogeneity may have hypothetically had an influence on the absolute and particularly relative reliability of the HRV measurement. As far as we are aware, this has not been discussed in the scientific field as yet. Just as in the previous findings, the results of the presented study indicate once again that the within-individual reliability of short-term HRV measurement, carried out by means of orthoclinostatic stimulation, is relatively low even if the study group was divided into the two more homogeneous subgroups.

The methods utilized to assess autonomic regulation require stationarity of HRV recordings. Non-stationarities are expected to be present in any HRV recording, even if experimental protocols are designed to steadily activate only the specific physiological mechanism under study and to keep external influences under control by careful supervision of the experimental setting (Magnini et al., 2011). Mood, alertness or mental activity changes are difficult to control and are probable reasons for the large random variation within individuals. The extremely similar SEM was observed for all HRV parameters between males and females, which do not allow us to make the conclusion of the different intra-session stability between these groups.

Relative reliability, which indicates how results correspond to real values (i.e. inter-individual reliability), is mostly around ICC = 0.8. Therefore, according to these between-individual differences, this HRV measurement and analysis may be considered sufficiently reliable or stable within the one session. As mentioned above, however, the most important of the reliability studies is SEM, not the correlation analysis. The estimated sample size needed to detect a significant mean change ($\geq$ 30 % of between-individual standard deviation) ranged from 5 to 32 for males and from 4 to 33 for females. The SEM magnitude is the reason for this diversity (Hopkins, 2000).

HRV magnitude influence on intra-session stability

The sorting of the study group according to HRV results, or more precisely according to cardiac autonomic activity, is another possible means of increasing its homogeneity. We could not proceed from any published research, because it has not been discussed in the scientific field as yet. We consequently chose the spectral parameter PT in the supine position as a marker for separating individuals with high or low HRV. Total Power is the sum of all the spectral component powers and is considered an indicator of total HRV (Task Force, 1996). The reliability level is consequently analysed separately for these two subgroups and compared with one another. Our pilot statistical procedure did not reveal the group difference if they were carried out according to the third or fifth supine interval. The sorting of the groups according to the PT in the first supine period is not appropriate because this interval serves the function of standardization.

The HRV directly decreases along with the slowing heart rate (HR) due to the shorter RR interval and the smaller HR oscillation. The HR increase has the opposite effect on the values of all the HRV parameters. There is consequently a need to be cautious in HRV change interpretation when the HR is different (Nieminen et al., 2007). The negative correlation of HR and HRV is similarly presented by Tsuji et al. (1996) or Kuch et al. (2001). We can also confirm that there are lower average HR values in the High PT group in comparison with the Low PT group.

It is apparent from Table 4 that the absolute reliability level in both subgroups is quite similar. Based on these presented results, it cannot be unequivocally concluded that there is a different intra-session stability level in individuals with a high or low HRV. The greatest difference (only one tenth of the ln value) is found in 95% LoA for PLF. The interval width for PLF, LF/HF and LFnu is much higher in the High PT group than the Low PT group. The considerable intra-individual variability in both subgroups must be, however, pronounced. Of interest is the fact that, according to the paired Student's t-test, there is not a test-retest zero mean change in the Low PT group (apart from the Mean RR) in the supine position, unlike all the HRV parameters in the High PT. The magnitude of the zero mean change expressed by the ES corresponds to this fact.

The correlation analysis reveals sufficient (ICC > 0.6) or even high (ICC > 0.8) relative reliability. The only exception is the PLF parameter in the High PT group in the supine position (ICC = 0.56). Significant differences between both subgroups cannot be seen.

As has been mentioned above, the SEM also directly influences the sample size needed to detect a significant mean change. Similarly to the previous part, the required sample size is not meaningfully different between the groups.

CONCLUSIONS

The research results are in accordance with previously published conclusions (Cipryan & Litschmannova, 2013), which present the low within-individual reliability of the HRV measurement. The study group heterogeneity may be one of the causes. The study group was therefore divided according to gender and results. Nevertheless, there are not any significant differences in the absolute or relative intra-session stability between these more homogeneous groups and we can conclude that HRV measurement reproducibility is not influenced by gender or HRV magnitude. It should be mentioned here that the reason for the large within-individual variability can be a measurement or analysis error as well as the natural physiological HRV oscillation. As the relative reliability is at least primarily sufficient, we would tend to support the second explanation.

REFERENCES

1. Antelmi I, De Paula, R.S., Shinzato A.R., Peres, C.A., Mansur A.J., & Grupi, C.J. (2004). Influence of age, gender, body mass index, and functional capacity on heart rate variability in a cohort of subjects without heart disease. *Am J Cardiol, 93(3),* pp.381-385.
2. Atkinson, G., & Nevill, A.M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sport medicine. *Sports Med, 26(4),* pp.217-238.
3. Bland, J.M., & Altman, D.G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet, 8(1),* pp.307-310.
4. Bland, J.M., & Altman, D.G. (1999). Measuring agreement in method comparison studies. *Stat Methods Med Res, 8(2),* pp.135-160.
5. Billman, G.E. (2002). Aerobic exercise conditioning: a nonpharmacological antiarrhythmic intervention. *J Appl Physiol, 92(2),* pp.446-454.
6. Bonnemeier, H., Wiegand, U.K.H., Brandes, A., Kluge, N., Katus, H.A., & Richardt, G. (2003). Circadian profile of cardiac autonomic nervous modulation in healthy subjects: Differing effects of aging and gender on heart rate variability. *J Cardiovasc Electr, 14(8),* pp.791-799.
7. Britton, A., Shipley, M., Malik, M., Hnatkova, K., Hemingway, H., & Marmot, M. (2007). Changes in heart rate and heart rate variability over time in middle-aged men and women in the general population (from the Whitehall II Cohort Study). *Am J Cardiol, 100(3),* pp.524-527.
8. Brook, R.D., & Julius, S. (2000). Autonomic imbalance, hypertension, and cardiovascular risk. *Am J Hypertens, 13(6)*, pp.112S-122S.
9. Carter, J.B., Banister, E.W., & Blaber, A.P. (2003). Effect of endurance exercise on autonomic control of heart rate. *Sports Med, 33(1),* pp.33-46.
10. Cipryan, L., & Litschmannova, M. (2013). Intra-day and inter-day reliability of heart rate variability measurement. *J Sport Sci, 31(2),* pp.150-158.
11. Davy, K.P., Miniclier, N.L., Taylor, J.A., Stevenson, E.T., & Seals, D.R. (1996). Elevated heart rate variability in physically active postmenopausal women: A cardiprotective effect? *Am J Physiol, 271(2 part 2),* pp. H455-H460.
12. Dekker, J.M., Crow, R.S., Folsom, A.R., Hannan, P.J., Liao, D., Swenne, C.A., & Schouten, E.G. (2000). Low heart rate variability in a 2-minute rhythm strip predicts risk of coronary heart disease and mortality from several causes: The ARIC Study. *Circulation, 102(11),* pp.1239-1244.
13. De Meersman, R.E., & Stein, P.K. (2000). Vagal modulation and aging. *Biol Psychol, 74(2),* pp.165-173.
14. Hedelin, R., Wiklund, U., Bjerle, P., & Henriksson-Larsén, K. (2000) Pre- and post-season heart rate variability in adolescent cross-country skiers. *Scand J Med Sci Sport, 10(5*), pp.298-303.

15. Hjortskov, N., Rissén, D., Blangsted, A.K., Fallentin, N., Lundberg, U., & Sogaard, K. (2004). The effect of mental stress on heart rate variability and blood pressure during computer work. *Eur J Appl Physiol, 92(1-2)*, pp.84-89.
16. Hopkins, W.G. (2000). Measures of reliability in sports medicine and science. *Sports Med, 30(1),* pp.1-15.
17. Hopkins, W.G. (2009). *Calculating the reliability intraclass correlation coefficient and its confidence limits (Excel spreadsheet).* newstats.org/xICC.xls.
18. Huikuri, H.V., Pikkujamsa, S.M., Airaksinen, K.E.J., Ikaheimo, M.J., Rantala, A.O., Kauma, H., Lilja, M., & Kesaniemi, Y.A. (1996). Sex-related differences in autonomic modulation of heart rate in middle-aged subjects. *Circulation, 94(2),* pp.122-125.
19. Jensen-Urstad, K., Storck, N., Bouvier, F., Ericson, M., Linbland, L.E., & Jensen-Urstad, M. (1997). Heart rate variability in healthy subjects is related to age and tender. *Acta Physiol Scand, 160(3),* pp.235-241.
20. Kobayashi, H. (2007). Inter- and intra-individual variations of heart rate variability in Japanese males. *J Physiol Anthropol, 26(2),* pp.173-177.
21. Koskinen, T., Kähönen, M., Jula, A., Mattsson, N., Laitinen, T., Keltikangas-Järvinen, L., Viikari, J., Välimäki, I., Rönnemaa, T., & Raitakari, O.T. (2009). Metabolic syndrome and short-term heart rate variability in young adults. *Diabetic Med, 26(4),* pp.354-361.
22. Kuch, B., Hense, H.W., Sinnreich, R., Kark, J.D., Von Eckardstein, A., Sapoznikov, D., & Bolte, H-D. (2001). Determinants of short-period heart rate variability in the general population. *Cardiology, 95(3),* pp.131-138.
23. Kuo, T.B.J., Lin, T., Yang, C.C.H., Li, C., Chen, C., & Chou, P. (1999). Effect of aging on gender differences in neural control of heart rate. *Am J Physiol, 277(6pt2),* pp.H2233–H2239.
24. Leicht, A.S., Hirning, D.A., & Allen, G.D. (2003). Heart rate variability and endogenous sex hormones during the menstrual cycle in young women. *Exp Physiol, 88(3),* pp.441-446.
25. Magagnin, V., Bassani, T., Bari, V., Turiel, M., Maestri, R., Pinna, G.D., & Porta, A. (2011). Non-stationarities significantly distort short-term spectral, symbolic and entropy heart rate variability indices. *Physiol Meas, 32(11),* pp.1775-1786.
26. Malpas, S.C. (2002). Neural influences on cardiovascular variability: possibilities and pitfalls. *Am J Physiol-Heart C, 282(1)*, pp.H6-H20.
27. Martinmäki, K., Rusko, H., Kooistra, L., Kettunen, J., & Saalasti S. (2006). Intraindividual validation of heart rate variability indexes to measure vagal effects on hearts. *Am J Physiol-Heart C, 290(2),* pp.H640-H647.
28. Migliaro, E.R., Contreras, P., Bech, S., Etxagibel, A., Castro, M., Ricca, R., & Vicente, K. (2001). Relative influence of age, resting heart rate and sedentary life style in short-term analysis of heart rate variability. *Braz J Med Biol Res, 34(4),* pp.493-500.
29. Nieminen, T., Kähönen, M., Kööbi, T., Nikus, K., & Viik J. (2007). Heart rate variability is dependent on the level of heart rate. *Am Heart J, 154(1),* pp.e13.
30. Pinna, G.D., Maestri, R., Torunski, A., Danilowicz-Szymanowicz, L., Szwoch, M., La Rovere, M.T., & Raczak, G. (2007). Heart rate variability measures: A fresh look at reliability. *Clin Sci, 113(3),* pp.131-140.
31. Ramaekers, D., Ector, H., Aubert, A.E., Rubens, A., & Van De Werf, F. (1998). Heart rate variability and heart rate in healthy volunteers. Is the female autonomic nervous system cardioprotective? *Eur Heart J, 19(9)*, pp.1334–1341.
32. Salinger, J., & Gwozdziewicz, M. (2008). *Systémy používané pro vyšetření krátkodobé variability srdeční frekvence. In: Javorka K, et al.* (eds) Variabilita frekvencie srdca: Mechanismy, hodnotenie, klinické využitie, OSVETA, Martin, pp 57-60.

33. Sandercock, G.R., Bromley, P.D., & Brodie, D.A. (2005). Effects of exercise on heart rate variability: inferences from meta-analysis. *Med Sci Sport Exer, 37(3),* pp.433-439.
34. Stein, P.K., Barzilay, J.I., Domitrovich, P.P., Chaves, P.M., Gottdiener, J.S., Heckbert, S.R., & Kronmal, R.A. (2007). The relationship of heart rate and heart rate variability to non-diabetic fasting glucose levels and the metabolic syndrome: The Cardiovascular Health Study. *Diabetic Med, 24(8),* pp.855-863.
35. Strauss, E., Sherman, E.M.S., & Spreen, O. (2006). *A compendium of neuropsychological test: administration, norms, and commentary*. New York: Oxford University Press.
36. Task Force of the European Society of Cardiology, North American Society of Pacing and Electrophysiology Heart rate variability: standards of measurement, physiological interpretation, and clinical use. *Circulation, 93(5),* pp.1043-65.
37. Tsuji, H., Venditti, F.J.J.R., Manders, E.S., Evans, J.C., Larson, M.G., Feldman, C.L., & Levy, D. (1996). Determinants of heart rate variability. *J Am Coll Cardiol, 28(6),* pp.1539-1546.
38. Umetani, K., Singer, D.H., & Mccraty, R. (1998). Atkinson M Twenty-four hour time domain heart rate variability and heart rate: relations to age and gender over nine decades. *J Am Coll Cardiol, 31(3),* pp.593-601.
39. Yamamoto, Y., Hughson, R.L., & Peterson, J.C. (1991). Autonomic control of heart rate during exercise studied by heart rate variability spectral analysis. *J Appl Physiol, 71(3),* pp.1136-1142.