Validation of Heart Rate Monitor Polar RS800 for Heart Rate Variability Analysis During Exercise

David Hernando,[1,2] Nuria Garatachea,[3,4,5] Rute Almeida,[6,1,2] Jose A. Casajús,[3,4,5] and Raquel Bailón[1,2]

[1]*Biomedical Signal Interpretation & Computational Simulation (BSICoS) Group, Aragon Institute of Engineering Research (I3A), IIS Aragón, University of Zaragoza, Spain;* [2]*Bioengineering, Biomaterials and Nanomedicine Networking Biomedical Research Center (CIBER-BBN), Zaragoza, Spain;* [3]*Physiatry and Nursing Department, Health and Sport Sciences Faculty, GENUD (Growth, Exercise, Nutrition and Development) research group. AgriFood Institute of Aragon IA2 (University of Zaragoza-CITA), Zaragoza, Spain;* [4]*Aragon Institute for Health Research (IIS Aragón);* [5]*Physiopathology of Obesity and Nutrition Networking Biomedical Research Center (CIBERObn); and* [6]*Mathematics Department, Sciences Faculty, University of Porto, (FCUP), Mathematics Center, University of Porto (CMUP), Porto, Portugal*

Abstract

Hernando, D, Garatachea, N, Almeida, R, Casajús, JA, and Bailón, R. Validation of heart rate monitor Polar RS800 for heart rate variability analysis during exercise. *J Strength Cond Res* 32(3): 716–725, 2018–Heart rate variability (HRV) analysis during exercise is an interesting noninvasive tool to measure the cardiovascular response to the stress of exercise. Wearable heart rate monitors are a comfortable option to measure interbeat (RR) intervals while doing physical activities. It is necessary to evaluate the agreement between HRV parameters derived from the RR series recorded by wearable devices and those derived from an electrocardiogram (ECG) during dynamic exercise of low to high intensity. Twenty-three male volunteers performed an exercise stress test on a cycle ergometer. Subjects wore a Polar RS800 device, whereas ECG was also recorded simultaneously to extract the reference RR intervals. A time–frequency spectral analysis was performed to extract the instantaneous mean heart rate (HRM), and the power of low-frequency (PLF) and high-frequency (PHF) components, the latter centered on the respiratory frequency. Analysis was done in intervals of different exercise intensity based on oxygen consumption. Linear correlation, reliability, and agreement were computed in each interval. The agreement between the RR series obtained from the Polar device and from the ECG is high throughout the whole test although the shorter the RR is, the more differences there are. Both methods are interchangeable when analyzing HRV at rest. At high exercise intensity, HRM and PLF still presented a high correlation ($\rho > 0.8$) and excellent reliability and agreement indices (above 0.9). However, the PHF measurements from the Polar showed reliability and agreement coefficients around 0.5 or lower when the level of the exercise increases (for levels of O_2 above 60%).

Key Words exercise stress test, wearable heart rate device, frequency domain analysis, reliability, agreement

Introduction

Heart rate variability (HRV) remains a powerful source for autonomic nervous system (ANS) assessment by using simple and noninvasive techniques. Spectral analysis of HRV at rest reveals at least 2 components: one low frequency component (LF) between 0.04 and 0.15 Hz and one high frequency component (HF), usually measured between 0.15 and 0.4 Hz. The HF component is related to the parasympathetic stimulation of the sinoatrial node and mainly because of respiratory sinus arrhythmia. The LF component is affected by both sympathetic and parasympathetic modulation and its interpretation is controversial. The ratio between the power in the LF and HF bands has been proposed to evaluate the sympathovagal balance controlling the heart rate (36). Analysis of HRV can be useful to diagnose and monitor several pathologies that are related to some ANS dysfunction. However, HRV monitoring is also interesting in subjects that do not present an apparent disease because it can be used to monitor the sleep or stress level, see Tobaldini et al. (31), Ranganathan et al. (26) and Kumar et al. (17).

In recent years, HRV analysis has gained increasing interest in sports and training sciences. Sports physiologists use HRV as a noninvasive measurement of autonomic

Address correspondence to Mr. David Hernando, dhernand@unizar.es.

32(3)/716–725

Journal of Strength and Conditioning Research

Table 1. Study population characteristics (23 male volunteers): age, height, mass, body mass index, and maximum oxygen consumption ($\dot{V}O_2$ max).*

Age (y)	34.8 ± 5.0
Height (cm)	178.4 ± 5.7
Mass (kg)	74.8 ± 7.8
Body mass index ($kg \cdot m^{-2}$)	23.5 ± 2.5
$\dot{V}O_2$ max (ml $O_2 \cdot kg^{-1} \cdot min^{-1}$)	57.8 ± 6.1

*Mean ± *SD*.

changes because of exercise training and the cardiovascular response to the stress of exercise, see Borresen et al. (8). Hottenrott et al. (15) show that significant differences can be observed in healthy subjects because of regular aerobic training: not only a reduction in mean heart rate both at rest and during submaximal exercise, but also an increase in autonomic efferent activity and a shift in favor of enhanced vagal modulation. Martinmäki et al. (22) also show that low-dose endurance training enhances vagal control during exercise, but does not alter resting vagal HR control. The study of HRV is also valuable for physicians during exercise. Analysis of HRV during stress testing achieved higher accuracy than any other stress electrocardiogram (ECG) indexes to discriminate ischemia, see Bailón et al. (6). Pradhapan et al. (25) have used HRV analysis during recovery after exercise to predict risk of mortality. However, HRV interpretation during exercise is still a matter of debate. It is necessary to take into account analysis methods, population characteristics, training levels, intensity, and duration of exercise. A multidisciplinary approach between cardiologists, exercise and pulmonary physiologists, coaches, and biomedical engineers would be desirable to interpret HRV during exercise, see Aubert et al. (2).

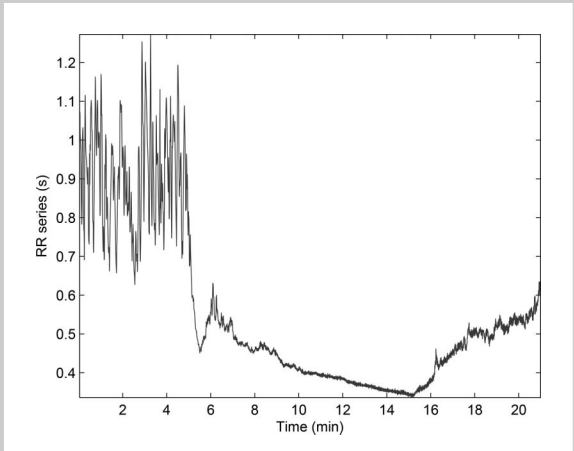

Figure 1. Example of interbeat (RR) series during the resting (0–5 minutes), exercise (5–15 minutes), and recovery (15-end) phases for one subject.

The analysis of HRV during exercise faces several challenges. First, it is particularly difficult to obtain reliable and robust QRS detections during exercise, mainly because of the significant levels of noise observed in this context (Llamedo et al. (19)) and changes in beat morphology (Drezner et al. (10)). Because of the nonstationary nature of HRV during exercise, time-varying spectral analysis should be used, such as parametric methods like autoregressive models with time-varying parameters, or nonparametric methods like quadratic time–frequency representations, the latter including the Wigner–Ville distribution and its filtered versions (Mainardi (20)). Moreover, respiratory rate increases with exercise intensity and can exceed the upper limit of the classical HF band (0.4 Hz), making it necessary to redefine the HF band, as shown in Bailón et al. (3). Mean heart rate also increases with exercise intensity and affects HRV parameters. A correction of HRV parameters with mean heart rate should be performed to separate changes in ANS stimulation from changes in mean heart rate, see Bailón et al. (5) and Sacha et al. (29,30). Finally, a component related to cardiolocomotor coupling has been observed in HRV during exercise. This component is centered at pedaling or running stride frequency and can mislead the interpretation of HRV as a marker of ANS modulation, see Bailón et al. (4).

Currently, there are several mobile and easy-to-use heart rate monitors that allow keeping a record of the interbeat intervals (RR intervals) during physical activity. Commercial devices, like Polar heart rate monitors, have been used by scientists for HRV analysis in sport sciences, medicine, and other fields of research: see Kaber et al. (16), Kumar et al. (17) and Turner et al. (32). Recent studies have validated Polar devices against different ECG systems. Works of Rezende Barbosa et al. (27), Gamelin et al. (12), and Giles et al. (13) show promising results and claim that, at rest, these devices are able to yield RR intervals series for HRV analysis as reliably as those obtained by ECG. Wallén et al. (33) studied the differences in the reliability according to the gender, claiming that the reliability between Polar and ECG measurements was weaker for women. Also, Weippert et al. (35) show that HRV measurements from Polar device are reliable during supine and sitting rest, walking, and moderate-to-vigorous static exercise of the limbs. To the authors' knowledge, no study has validated heart rate monitor Polar RS800 during dynamic exercise of high intensity.

The aim of this study is to evaluate the agreement and reliability between HRV analysis derived from RR series recorded by HR monitor Polar RS800 and HRV analysis

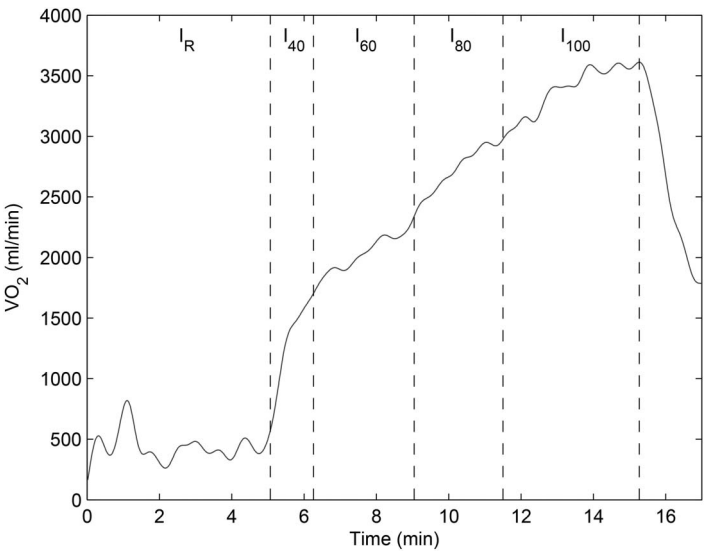

Figure 2. Example of oxygen consumption ($\dot{V}O_2$) signal and intervals I_R, I_{40}, I_{60}, I_{80} and I_{100} (resting phase, 0–40%, 40–60%, 60–80% and 80–100% of $\dot{V}O_2$, respectively) for one subject.

derived from a simultaneous ECG recording during dynamic exercise of low, medium, and high intensity.

Methods

Experimental Approach to the Problem

This study validates a Polar RS800 device in 23 healthy male volunteers during an exercise test. A high-resolution ECG is simultaneously recorded to extract the RR intervals and use them as a reference. A time–frequency spectral analysis is performed to extract the mean heart rate (HRM), the power of low frequency (PLF) and the power of high frequency (PHF), the latter centered on the respiratory frequency. A Bland–Altman plot is used with the RR intervals derived from the ECG and the Polar device to graphically see the discrepancy between both measurements across the whole range. Reliability and agreement coefficients are also computed in different intervals related to the level of oxygen consumption for each subject.

Subjects

A total of 23 volunteers agreed to participate in the study. All of them were apparently healthy subjects and regularly participate in sports activities. Written informed consent was obtained from each subject. The study protocol was approved by the University of Zaragoza and was in accordance with the Declaration of Helsinki for Human Research of 1974 (last modified in 2008). Table 1 shows study population characteristics.

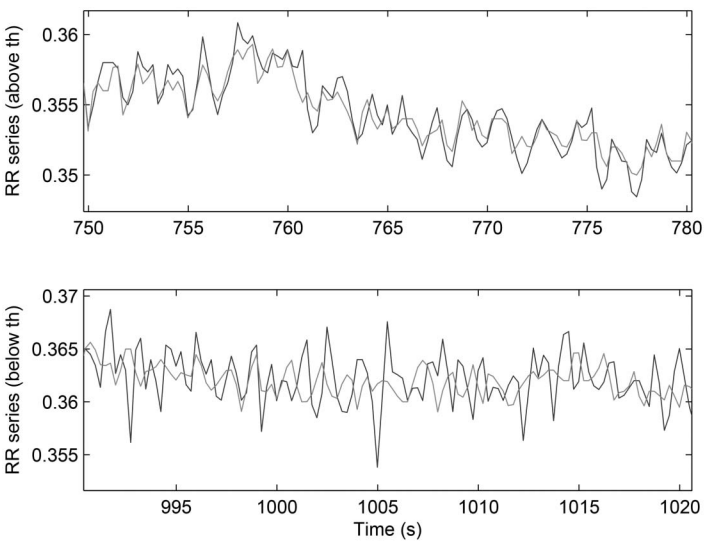

Figure 3. Example of an interval of the interbeat (RR) series with a high correlation coefficient $\rho = 0.89$ (upper panel) and low correlation coefficient $\rho = 0.65$ (lower panel): RR_{ECG} in blue, RR_{POL} in red.

Experimental Protocol. All the subjects completed a submaximal test on a cycle ergometer. Before the tests, they were asked to adhere to the instructions described in the study of Wasserman et al. (34): (a) wear comfortable, loose-fitting clothing; (b) drink plenty of fluids over the 24-hour period preceding the test; (c) avoid food, tobacco, alcohol, and caffeine for 3 hours before taking the test; (d) avoid exercise or strenuous physical activity on the day of the test; and (e) get an adequate amount of sleep (6–8 hours) the night before the test.

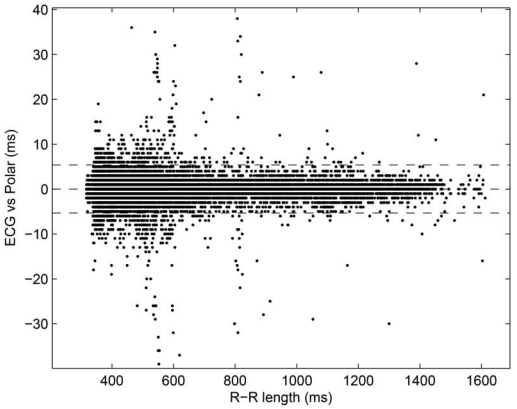

Figure 4. Bland–Altman plot: RR_{ECG} vs. RR_{POL}. Mean of the difference of the RR measurements ±1.96*std values (limits of agreement, LOA).

The exercise test was divided into 2 different phases: resting and exercise phase. During the resting phase, the subjects were continuously monitored while seated at rest for 5 minutes, without any movement or talking, to measure resting cardiorespiratory variables. The exercise phase started on the cycle ergometer at 75 W work load, increasing at a rate of 25 $W \cdot min^{-1}$. The cadence frequency was fixed at 80 rpm. This phase lasted until the subject reached his 90% maximum heart rate, which was determined in previous tests by a physician. Then, the work load was kept constant for 2 more minutes. The recovery phase consisted of 5 minutes of pedaling at free cadence, but it is not included in this study because of its heterogeneity among the subjects. Figure 1 shows an example of the evolution of the RR series throughout the test.

The test was divided into 5 different intervals. In addition to the resting interval (I_R), which is associated with the resting phase (5 minutes before the exercise), the oxygen consumption ($\dot{V}O_2$) signal was used to establish 4 intervals during the exercise phase: 0–40%, 40–60%, 60–80%, and 80–100% of the O_2 consumption variation during exercise for each subject (Figure 2). The basal value of $\dot{V}O_2$ is obtained as the mean value during the resting phase, whereas the maximum value is found at the peak of the $\dot{V}O_2$ signal. These intervals are denoted as I_{40}, I_{60}, I_{80}, and I_{100}, respectively. Note that each interval has different length among the subjects.

Procedures

Data Acquisition and Preprocessing. The ECG signal, RR intervals, respiratory frequency, and $\dot{V}O_2$ were recorded throughout the 3 phases of the test.

Ventilatory and exchange gases were analyzed breath-by-breath by an open-circuit sampling system (Oxycon Pro; Jaeger-Viasys Healthcare, Hoechberg, Germany). The metabolic cart was calibrated with a known gas mixture (16% oxygen, O_2, and 5% carbon dioxide, CO_2) and volume before the first test each day as recommended by the company. Both respiratory frequency and $\dot{V}O_2$ were interpolated at 4 Hz and low-pass filtered with a cut-off frequency of 0.01 Hz.

Interbeat intervals were recorded beat-to-beat using an HR monitor (RS800; Polar Electro Oy, Kempele, Finland), which uses a sampling frequency of 1,000 Hz for the ECG signal. Moreover, the ECG was simultaneously recorded using a high-resolution Holter (Mortara 48-hour H12+; Mortara Instrument, Milwaukee, WI, USA) with a sampling frequency of 1,000 Hz. For each subject, the QRS detection marks were extracted from the ECG using a multilead approach by a wavelet-based detector described in Martínez et al. (23) with optimized parameters for noisy environments described in Hernando et al. (14), and each detection was manually verified by an operator with a dedicated interface. Interbeat intervals from the ECG (RR_{ECG}) were obtained as the difference of each consecutive beat occurrences. Interbeat intervals from Polar (RR_{POL}) were directly obtained from the device. The delay between RR_{ECG} and RR_{POL} was estimated as that lag which maximizes their cross-correlation. Subsequently, the 2 series were synchronized by correcting this delay.

Table 2. Bias, limits of agreement (LOA), and percentage of paired interbeat (RR) measurements out of the LOA.*

RR (ms)	BIAS (ms)	LOA (ms)	Out LOA (%)
I_{100}: 359.7 ± 42.7	−0.0192	−5.198, 5.160	4.80
I_{180}: 435.9 ± 33.5	−0.0022	−4.234, 4.229	3.99
I_{60}: 507.9 ± 38.4	−0.0737	−6.075, 5.927	3.81
I_{40}: 635.5 ± 89.1	0.1484	−5.981, 6.277	3.11
I_R: 1,167.3 ± 442.7	0.0616	−5.174, 5.298	1.67
Whole range	0.0205	−5.341, 5.382	3.32

*Different intervals are based on the oxygen consumption ($\dot{V}O_2$): I_R, I_{40}, I_{60}, I_{80}, and I_{100} for the resting phase, 0–40%, 40–60%, 60–80%, and 80–100% of $\dot{V}O_2$, respectively.

Heart Rate Variability Indices. The instantaneous heart rate signal, $d_{HR}(n)$, is derived from both RR intervals series, following a method based on the integral pulse frequency modulation (IPFM) model, which also accounts for the presence of ectopic beats (Mateo et al. (24)), and sampled at 4 Hz. This signal is high-pass-filtered to remove the mean heart rate $d_{HRM}(n)$ (very low-frequency components, up to 0.03 Hz) and it

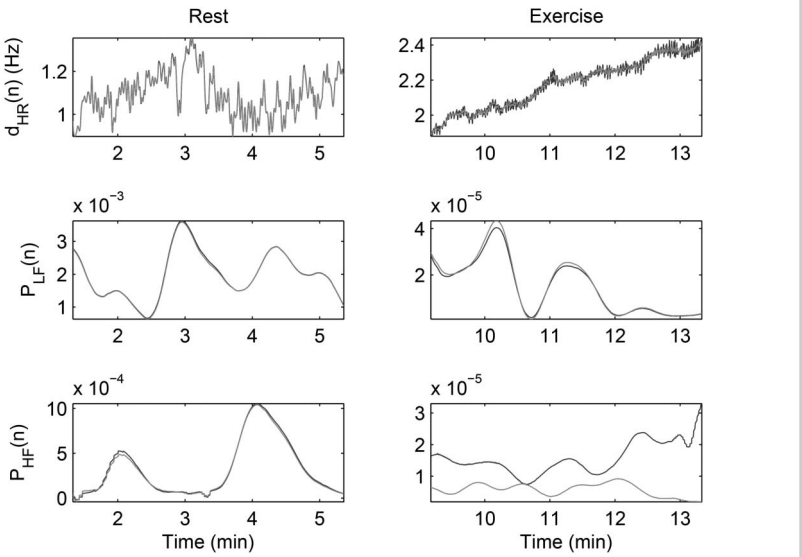

Figure 5. Example of d_{HR} (n), P_{LF} (n) and P_{HF} (n): Resting phase (left) and exercise phase (right). Blue denotes signals derived from ECG and red denotes signals derived from Polar. Note that the axis has different scales.

is also corrected by it: $m(n) = \frac{(d_{HR}(n) - d_{HRM}(n))}{d_{HRM}(n)}$, see Bailón et al. (5).

The smoothed pseudo Wigner–Ville distribution (SPWVD) was applied to m(n) to estimate the time-varying spectral properties of both HRV signals (Martin et al. (21)). This time–frequency analysis allows to study the evolution of the power of the frequency components through time. To suppress the interference terms, both time and frequency smoothing windows were chosen to be Hamming windows of length $2N-1 = 203$ (about 50 seconds) and $2K-1 = 513$ samples (about 128 seconds), respectively, as described in Bailón et al. (4). The instantaneous power in the low- and high-frequency band, $P_{LF}(n)$ and $P_{HF}(n)$ respectively, was extracted throughout the entire exercise test. Low-frequency band ranged from 0.04 to 0.15 Hz. The high-frequency band was centered on the respiratory frequency with a bandwidth of 0.25 Hz. The lower limit of the HF band was never below 0.15 Hz, and the upper limit was never above the half of mean heart rate (Bailón et al. (4)).

Statistical Analyses

Reliability and Agreement Analysis. To study the reliability and agreement between HRV derived from the Polar records and from the ECG signal, 4 steps were proposed:

a. A Bland–Altman plot (Bland et al. (7)) was used to visualize both RR_{ECG} and RR_{POL} series. The bias, the limits of agreement (LOA), and the percentage of paired RR measurements that are out of the LOA were computed in each interval.
b. Pearson's correlation coefficient (Acton (1)), ρ, was used to quantify the linear relationship between the ECG and Polar following signals:
   - Mean heart rate: d_{HRM} (n)
   - Low-frequency power: P_{LF} (n)
   - High-frequency power: P_{HF} (n)

To determine if 2 signals present a high linear correlation, a threshold for the correlation coefficient ρ was set to 0.8. Figure 3 shows in the upper panel a case of RR series from a subject that presents a ρ value above the threshold (RR_{ECG} in blue, RR_{POL} in red), whereas lower panel shows a case with ρ value below the threshold.

For each interval (I_R, I_{40}, I_{60}, I_{80}, and I_{100}), a mean value was obtained (HRM, PLF, and PHF from d_{HRM} (n), P_{LF} (n), and P_{HF} (n) respectively) and the correlation coefficient was evaluated for every subject. Also, in those intervals with a ρ value above the threshold, a relative error was computed as the mean of the difference between the 2 signals (in absolute value), normalized by the mean value in the ECG signal.

Table 3. Number of subjects (out of 23) with a significant $\rho > 0.8$ and the relative error (in brackets).*

	I_R (%)	I_{40} (%)	I_{60} (%)	I_{80} (%)	I_{100} (%)
HRM	23 (0.58)	23 (0.09)	23 (0.05)	23 (0.06)	23 (0.07)
PLF	19 (3.82)	23 (2.59)	23 (6.01)	22 (5.49)	22 (12.73)
PHF	22 (9.26)	20 (15.77)	17 (29.93)	16 (63.91)	12 (90.25)

*Different intervals are based on the oxygen consumption ($\dot{V}O_2$): I_R, I_{40}, I_{60}, I_{80}, and I_{100} for the resting phase, 0–40%, 40–60%, 60–80%, and 80–100% of $\dot{V}O_2$, respectively.

Table 4. Reliability (CCC, ICC) and agreement (A) coefficients for each interval, with the confidence intervals.*

	I_R	I_{40}	I_{60}	I_{80}	I_{100}
RR					
CCC	0.9776 (0.9679–0.9898)	1.0000 (0.9999–1.0000)	0.9999 (0.9998–1.0000)	1.0000 (1.0000–1.0000)	1.0000 (1.0000–1.0000)
ICC	0.9994 (0.9983–0.9998)	1.0000 (0.9999–1.0000)	0.9999 (0.9998–1.0000)	1.0000 (1.0000–1.0000)	1.0000 (1.0000–1.0000)
A	0.9937 (0.9868–0.9971)	0.9994 (0.9982–0.9998)	0.9989 (0.9959–0.9998)	0.9995 (0.9992–0.9997)	0.9993 (0.9987–0.9997)
HRM					
CCC	0.9989 (0.9979–0.9998)	0.9999 (0.9998–1.0000)	1.0000 (1.0000–1.0000)	1.0000 (1.0000–1.0000)	1.0000 (1.0000–1.0000)
ICC	0.9995 (0.9983–0.9998)	0.9999 (0.9998–1.0000)	1.0000 (1.0000–1.0000)	1.0000 (1.0000–1.0000)	1.0000 (1.0000–1.0000)
A	0.9938 (0.9870–0.9970)	0.9991 (0.9968–0.9998)	0.9995 (0.9990–0.9998)	0.9995 (0.9992–0.9997)	0.9995 (0.9990–0.9997)
PLF					
CCC	0.9677 (0.9404–0.9951)	0.9990 (0.9982–0.9999)	0.9951 (0.9929–0.9973)	0.9996 (0.9992–0.9999)	0.9996 (0.9995–0.9997)
ICC	0.9843 (0.9630, 0.9933)	0.9995 (0.9989, 0.9998)	0.9976 (0.9945, 0.9990)	0.9998 (0.9995, 0.9999)	0.9998 (0.9995–0.9999)
A	0.9222 (0.8685–0.9590)	0.9797 (0.9691–0.9868)	0.9562 (0.8790–0.9780)	0.9577 (0.9372–0.9706)	0.9210 (0.8672–0.9251)
PHF					
CCC	0.9674 (0.9528–0.9819)	0.9854 (0.9729–0.9978)	0.9296 (0.8740–0.9851)	0.8611 (0.7582–0.9640)	**0.3342 (0.0072–0.6756)**
ICC	0.9840 (0.9583–0.9935)	0.9929 (0.9817–0.9971)	0.9650 (0.9175–0.9852)	0.9283 (0.8319–0.9695)	**0.5121 (0.1860–0.7957)**
A	0.9069 (0.8565–0.9423)	0.7859 (0.6823–0.8564)	0.7160 (0.6390–0.7885)	**0.5641 (0.4504–0.6762)**	**0.5633 (0.4508–0.6687)**

*Values lower than 0.7 are bold. Different intervals are based on the oxygen consumption ($\dot{V}O_2$): I_R, I_{40}, I_{60}, I_{80}, and I_{100} for the resting phase, 0–40%, 40–60%, 60–80%, and 80–100% of $\dot{V}O_2$, respectively.

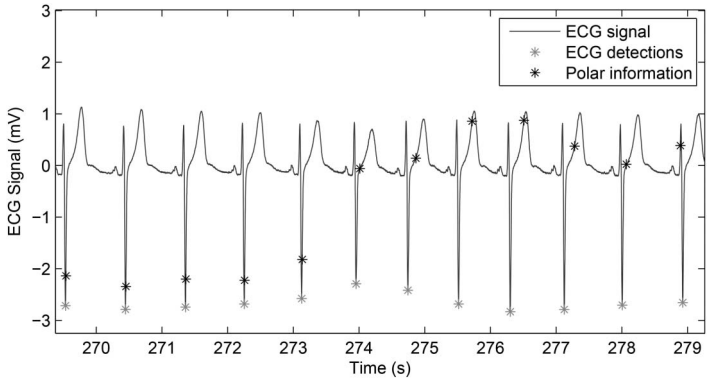

Figure 6. Example of ECG at rest with QRS detections (red marks) and Polar time occurrences (black marks).

c. To measure the interchangeability between measures, 2 reliability indexes were used: Lin's concordance correlation coefficient (CCC) and intraclass correlation coefficient (ICC).

The CCC index measures the reliability between 2 methods, and it determines how far the observed data deviate from the line of perfect concordance line at 45° on a square axis scatter plot, see Lin (18). Alternatively, ICC index that also measures reliability represents the ratio of between-sample variance and the total variance (between-sample and within-sample) to measure precision under the model of equal marginal distributions, see Fisher (11).

d. The agreement was measured by an information-based measure of disagreement (IBMD) proposed in Costa-Santos (9). This measurement is based on Shannons entropy and equals 0 when the observers agree (no disagreement: $xi = yi$), i.e., there is no information in the differences between methods X and Y. This measurement increases toward to 1 if the amount of information in the difference increases. The agreement (A) can be quantified as $A = 1 - \text{IBMD}$.

Results

Bland–Altman Plot

Figure 4 shows a Bland–Altman plot that evaluates the intermethod discrepancies between ECG and Polar RR measurements and the stability across a wider value range. The central and the upper and lower horizontal lines show the bias (mean) and the LOA (±1.96*std values), respectively, of the differences between both methods. A total of 27,552 paired RR measurements were used, from which 96.68% were contained within the range. For shorter RR, i.e., higher heart rates, the discrepancies are larger. Table 2 shows the bias, LOA, and the percentage of paired RR measurements out of the LOA for each interval and the whole data combined. The mean RR and standard deviation associated with each interval is also shown.

Pearson's Correlation Coefficient

Figure 5 shows an example of d_{HR} (n), P_{LF} (n), and P_{HF} (n) for one subject. Because of the differences in the axes, the resting and exercise phase are represented in different plots. Left panels show the evolution of these signals during the 5 minutes before the exercise (resting interval), with blue representing the signal derived from RR_{ECG} (n) and red the signal derived from RR_{POL} (n). Right panels show the exercise phase.

Table 3 shows how many subjects present a Pearson's coefficient above 0.8 in each interval for the different signals. Also, the error calculated between the ECG and Polar derived measurements are displayed in brackets.

Reliability and Agreement Coefficients (CCC, ICC, and A)

Table 4 shows the CCC and ICC values in each interval. Coefficients are lower in I_R than in I_{40}, which was not expected, but still excellent reliability, i.e., above 0.9. These coefficients decrease in the last intervals (near the exercise peak) for PHF. We consider coefficient values lower than 0.7 as markers of poor reliability. Table 4 also shows the agreement between both methods in each interval. It is very similar as the reliability coefficients, with decreasing values for PHF near the peak of exercise.

Discussion

In this study, HRV analysis derived from RR series recorded by HR monitor Polar RS800 has been validated against HRV analysis derived from a simultaneous ECG recording during dynamic exercise of low, medium, and high intensity. Analysis during the resting phase is in agreement with those from previous studies, which support the validity of HRV analysis from Polar at rest: Rezende et al. (27) reported high correlation coefficients between all spectral HRV parameters (LF, HF, and ratio LF/HF). Gamelin et al. (12) presented very similar results during the resting phase, with a correlation coefficient of $\rho > 0.99$. In a recent study, Giles et al. (13) validated the HRV parameters from Polar V800 at rest and showed that a strong correlation (ICC > 0.999) and a narrower LOA could be achieved, when using an extra correction layer to identify errors in the Polar signal and using the same software package for both the Polar and ECG signals. We have adopted a similar approach applying the same algorithms to RR series recorded by Polar and those derived from the ECG, but no extra correction layer was used in Polar annotations.

In this database, only 19 out of 23 recordings presented high correlation coefficient in PLF at rest, $\rho > 0.8$, which is unexpected. The reason may be that in the other 4 recordings, changes in QRS amplitude or morphology may confuse Polar detection algorithm, as displayed in Figure 6. These changes in beat detection point introduce additional variation in the Polar RR-intervals that are not present in the ECG annotations, justifying the lower linear correlation in PLF observed in these subjects during the resting phase. Nevertheless, those were not sufficient to disrupt the excellent reliability and agreement found, as can be seen in Table 4.

Regarding the exercise phase, the first issue to study is the agreement between the detected R-peaks, which is addressed assessing the agreement between RR series. It can be observed that for higher intensity levels, related to shorter RR values, the more differences there are between the RR series obtained from Polar device and from the ECG. The Bland–Altman plot shows how the discrepancy increases from 1.67% of the RR pairs out of the limits of agreement (at rest) to 4.80% in the last interval (I_{100}).

Regarding HRM at exercise, it shows an excellent correlation, reliability, and agreement during the whole test between both measurements. Higher loads of exercise do not decrease the performance of the Polar measurements. Power of low frequency also shows excellent correlation, reliability, and agreement during the whole exercise test, meaning that Polar device can follow heart rate oscillations up to 0.15 Hz and the ECG reference. Interestingly, the beat detection errors that occasionally happened at rest disappear in the exercise phase. These results support the use of these devices in those applications where only changes of slow oscillations (HRM and PLF) in heart rate matter while doing physical activities (sports, 24-hour Holter recordings…).

In this work, the relative error between PHF derived from Polar and that derived from the ECG is higher than for PLF in all the intervals considered, including the resting phase. Moreover, the disagreement in PHF increased with the increase in exercise intensity. Power in the high-frequency band is mainly related to respiratory sinus arrhythmia. During exercise, both respiratory frequency and depth change. Besides, in the HF band, there could be another components related to the pedal frequency (cardiolocomotor coupling), which are not related to the autonomic control, see Bailón et al. (4). Because a high level of exercise leads to a higher noise level, errors in beat detection are more likely to occur. A possible reason for this worse agreement is that Polar device uses some kind of filtering or prediction when a beat cannot be accurately detected, leading to an underestimation of HF power by Polar. However, we must say that in some recordings (12 out of 23), PHF measurements obtained from Polar coincided with those from the ECG even at the highest intensity interval ($\rho > 0.8$), so the correlation between HRV derived from Polar device and from the ECG strongly depends on subject characteristics and exercise performance.

Weippert et al. (35) validated HRV measurements from Polar device during walking and moderate-to-vigorous static exercise of the limbs with good results in both PLF and PHF. Unlike static exercise, high intensity dynamic exercise may contaminate the ECG signal with such amount of muscular noise and movement interferences that induces errors in the Polar detection algorithm. Although this may also happen in the reference ECG, we have used a QRS detection algorithm optimized for exercise conditions and all the detections have been visually checked.

In addition to the LF, HF and cardiolocomotor coupling components (the latter not reported in this work), a new frequency component was found around 1 Hz in HRV signals provided by the Polar device. This component did not appear in the HRV signal derived from the reference ECG, and it was present only in 8 subjects during the exercise phase. It is still unclear neither the origin of this component, nor the reason why it only appears in some recordings, but it should be taken into account when analyzing the total power of HRV obtained from Polar.

In conclusion, both methods are interchangeable when measuring HRM and PLF, regardless of the level of exercise (reliability and agreement > 0.9). However, the performance of the PHF measurements from Polar device decreases as the level of the exercise increases, with reliability and agreement coefficients around 0.5 or lower.

One potential limitation of this study is that only men are studied. Works like Wallén et al. (33) state that the reliability between Polar and ECG measurements is weaker for women; hence, a further study should validate the Polar device during exercise in women. Moreover, subjects included in the study regularly participate in sports activities ($\dot{V}O_2$ max: 57.8 ± 6.1 ml $O_2 \cdot kg^{-1} \cdot min^{-1}$); therefore, they are not representative of the general population, which may be a potential limitation related to practical application.

Practical Applications

This study validates a Polar heart rate monitor during exercise. The use of these wearables devices has recently grown, especially in casual sport. Heart rate variability indexes have normally been studied to measure chronic adaptation in athletes, mainly at rest, see Borresen et al. (8), Hottenrott et al. (15). It would also be interesting to record the RR series during the exercise and the recovery phases, which could provide more information regarding the autonomic balance behavior.

Using a Polar device, any athlete could have a recording of every training session and, later, use it to adjust the training load or evaluate the overtraining. In the clinic, exercise therapy is used to improve the HRV in patients with chronic heart failure (CHF), myocardial infarction (MI), coronary artery disease and hypertension, among others. Heart rate

variability parameters measured at frequency components up to 0.15 Hz (LF and lower) have been found to be predictors of mortality in patients with CHF or MI when being analyzed after the exercise sessions, see Routledge et al. (28). Therefore, being able to measure those HRV parameters also during the exercise could add new diagnostic information; as the case of ischemic patients, who show different HRV behavior during exercise depending on the ischemia etiology.

However, clinicians, coaches, and sports physiologists need to be cautious when interpreting the high-frequency HRV parameters provided by Polar during training. Results of this study suggest that RR series provided by Polar device are useful to study the evolution of slow oscillations in HRV, such as the changes in mean HR or the LF component.

Acknowledgments

This work is supported by the Diputación General de Aragón (DGA), Spain, through a fellowship with reference B195/12, by the Ministerio de Economía y Competitividad and FEDER (EU), under project TIN2014-53567-R and TEC2013-42140-R, by CIBER in Bioengineering, Biomaterials & Nanomedicine (CIBER-BBN), by Grupo Consolidado BSICoS ref:T96 from DGA, by Aragón Institute of Engineering Research (I3A), IIS Aragón and European Social Fund (EU). Research was partially supported by CMUP (UID/MAT/00144/2013), funded by FCT (Portugal) with national (MEC) and European structural funds through FEDER, under the partnership agreement PT2020. The computation was performed by the ICTS NANBIOSIS, more specifically by the High Performance Computing Unit of the CIBER-BBN at the University of Zaragoza. The authors also want to acknowledge the support from Mortara Instr Inc. The results of the current study do not constitute endorsement of the product by the authors or the journal.

References

1. Acton, F. *Analysis of Straight-Line Data*. New York, Dover Publications, 1966.
2. Aubert, A, Seps, B, and Beckers, F. Heart rate variability in athletes. *Sports Med* 33: 889–919, 2003.
3. Bailón, R, Laguna, P, Mainardi, L, and Sörnmo, L. Analysis of heart rate variability using time-varying frequency bands based on respiratory frequency. *Conf Proc IEEE Eng Med Biol Soc* 6674–6677, 2007.
4. Bailón, R, Garatachea, N, de la Iglesia, I, Casajús, J, and Laguna, P. Influence of running stride frequency in heart rate variability analysis during treadmill exercise testing. *IEEE Trans Biomed Eng* 60: 1796–1805, 2013.
5. Bailón, R, Laouini, G, Grao, C, Orini, M, Laguna, P, and Meste, O. The integral pulse frequency modulation with time-varying threshold: Application to heart rate variability analysis during exercise stress testing. *IEEE Trans Biomed Eng* 58: 642–652, 2011.
6. Bailón, R, Serrano, P, and Laguna, P. Influence of time-varying mean heart rate in coronary artery disease diagnostic performance of heart rate variability indices from exercise stress testing. *J Electrocardiol* 44: 445–452, 2011.
7. Bland, J and Altman, D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 8: 307–310, 1986.
8. Borresen, J and Lambert, M. Autonomic control of heart rate during and after exercise: Measurements and implications for monitoring training status. *Sports Med* 38: 633–646, 2008.
9. Costa-Santos, C, Antunes, L, Souto, A, and Bernardes, J. Assessment of disagreement: A new information-based approach. *Ann Epidemiol* 20: 555–561, 2010.
10. Drezner, JA, Fischbach, P, Froelicher, V, Marek, J, Pelliccia, A, Prutkin, JM, Schmied, CM, Sharma, S, Wilson, MG, Ackerman, MJ, Anderson, J, Ashley, E, Asplund, CA, Baggish, AL, Brjesson, M, Cannon, BC, Corrado, D, DiFiori, JP, Harmon, KG, Heidbuchel, H, Owens, DS, Paul, S, Salerno, JC, Stein, R, and Vetter, VL. Normal electrocardiographic findings: Recognising physiological adaptations in athletes. *Br J Sports Med* 47: 125–136, 2013.
11. Fisher, R. *Statistical Methods for Research Workers*. Edinburgh, Scotland: Oliver & Boyd, 1925.
12. Gamelin, F, Baquet, G, Berthoin, S, and Bosquet, L. Validity of the polar S810 to measure R-R intervals in children. *Int J Sports Med* 29: 134–138, 2008.
13. Giles, D and Draper, N. NeilW. Validity of the polar V800 heart rate monitor to measure RR intervals at rest. *Eur J Appl Physiol* 116: 563–571, 2016.
14. Hernando, D, Bailón, R, Almeida, R, and Hernández, A. QRS detection optimization in stress test recordings using evolutionary algorithms. *XLI Int Conf Comput Cardiol*: 737–740, 2014.
15. Hottenrott, K, Hoos, O, and Esperer, H. Heart rate variability and physical exercise. Current status. *Herz* 31: 544–552, 2006.
16. Kaber, D, Perry, C, Segall, N, and Sheik-Nainar, M. Workload state classification with automation during simulated air traffic control. *Int J Aviat Psychol* 17: 371–390, 2007.
17. Kumar, M, Weippert, M, Vilbrandt, R, Kreuzfeld, S, and Stoll, R. Fuzzy evaluation of heart rate signals for mental stress assessment. *IEEE Trans Fuzzy Syst* 15: 791–808, 2007.
18. Lin, L. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45: 255–268, 1989.
19. Llamedo, M and Martnez, JP. QRS detectors performance comparison in public databases. *XLI Int Conf Comput Cardiol*: 357–360, 2014.
20. Mainardi, L. On the quantification of heart rate variability spectral parameters using time–frequency and time-varying methods. *Philos Trans A Math Phys Eng Sci* 367: 255–275, 2009.
21. Martin, W and Flandrin, P. Wigner-Ville spectral analysis of nonstationary processes. *IEEE Trans Acoust Speech Signal Process* 33: 1461–1470, 1985.
22. Martinmäki, K, Häkkinen, K, Mikkola, J, and Rusko, H. Effect of low-dose endurance training on heart rate variability at rest and during an incremental maximal exercise test. *Eur J Appl Physiol* 104: 541–548, 2008.
23. Martínez, JP, Almeida, R, Olmos, S, Rocha, AP, and Laguna, P. A wavelet-based ECG delineator: Evaluation on standard databases. *IEEE Trans Biomed Eng* 51: 570–581, 2004.
24. Mateo, J and Laguna, P. Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal. *IEEE Trans Biomed Eng* 50: 334–343, 2003.
25. Pradhapan, P, Tarvainen, M, Nieminen, T, Lehtinen, R, Nikus, K, Lehtimäki, T, Kähönen, M, and Viik, J. Effect of heart rate correction on pre- and post-exercise heart rate variability to predict risk of mortality-an experimental study on the FINCAVAS cohort. *Front Physiol* 2014. doi: 10.3389/fphys.2014.00208.
26. Ranganathan, G, Rangarajan, R, and Bindhu, V. Evaluation of ECG signals for mental stress assessment using fuzzy technique. *Int J Secur Softw Eng* 1: 195–201, 2011.

27. Rezende Barbosa, M, Silva, N, Azevedo, F, Pastre, C, and Vanderlei, L. Comparison of Polar R RS800G3TM heart rate monitor with Polar R S810iTM and electrocardiogram to obtain the series of RR intervals and analysis of heart rate variability at rest. *Clin Physiol Funct Imaging*, 2014. doi: 10.1111/cpf.12203.

28. Routledge, FS, Campbell, TS, McFetridge-Durdle, JA, and Bacon, SL. Improvements in heart rate variability with exercise therapy. *Can J Cardiol* 26: 303–312, 2010.

29. Sacha, J, Sobon, J, Sacha, K, and Barabach, S. Heart rate impact on the reproducibility of heart rate variability analysis. *Int J Cardiol* 19: 207–216, 2013.

30. Sacha, J. Interaction between heart rate and heart rate variability. *Ann Noninvasive Electrocardiol* 168: 4257–4259, 2014.

31. Tobaldini, E, Nobili, L, Strada, S, Casali, K, Braghiroli, A, and Montano, N. Heart rate variability in normal and pathological sleep. *Front Physiol* 4: 294, 2013.

32. Turner, S, Eastwood, P, Cecins, N, Hillman, D, and Jenkins, S. Physiologic responses to incremental and self-paced exercise in COPD: A comparison of three tests. *Chest* 126: 766–773, 2004.

33. Wallén, M, 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. Wasserman, K. *Principles of Exercise Testing and Interpretation: Including Pathophysiology and Clinical Applications* (5th ed.). Philadelphia, Lippincott Williams & Wilkins, 2011.

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. Working group of ESC. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. *Eur Heart J* 17: 354–381, 1996.