

Validation of a smart shirt for heart rate variability measurements at rest and during exercise

Peter Schams¹  | Boris Feodoroff¹  | Jonas Zacher² | Angi Eibl² | Ingo Froböse¹

¹Institute of Movement Therapy and Movement-oriented Prevention and Rehabilitation Sciences, German Sport University, Cologne, Germany

²Institute of Cardiology and Sports Medicine; German Sport University, Cologne, Germany

Correspondence

Boris Feodoroff, Institute of Movement Therapy and Movement-oriented Prevention and Rehabilitation Sciences, German Sport University, 50933 Cologne, Germany.
Email: B.Feodoroff@dshs-koeln.de

Funding information

Bundesministerium für Wirtschaft und Energie, Grant/Award Number: ZF 4423401AK7

Abstract

Heart rate variability (HRV) monitoring is a promising option to estimate the autonomic nervous system regulation responding to exercise. Textiles with embedded sensors recording heartbeat intervals are a simple tool for data collection. The so-called smart shirts offer comfort for daily use and are managed easily. Their measurement accuracy for HRV calculation at rest is promising, but remains questionable during exercise. Therefore, the present study validated the Ambiotex smart shirt using HRV indices (root mean square of successive differences, rel. HF power [high-frequency power percentage of total power] and rel. LF [low-frequency power percentage of total power] power) during exercise. Eighty-three healthy participants (31 ± 6 years; 39 females, 44 males) completed an incremental exercise test on a bicycle ergometer wearing the smart shirt and an electrocardiogram simultaneously. We compared HRV indices of segments at rest (5 min), at warm-up (3 min) and twice at the exercise test (each 5 min). At rest and at warm-up, we observed excellent linear relationship ($r > 0.96$; $R^2 > 0.94$), excellent relative reliability (intraclass correlation coefficient ≥ 0.98 ; $\alpha \geq 0.98$) and acceptable agreement (bias $< 10\%$). During the exercise test, measurement accuracy declined with increasing intensity but remained high (> 0.8), although results for partial HRV indices were insufficient. In addition, percentage bias was unacceptable during an exercise test. However, the findings support the validity of the smart shirt for measuring HRV, especially at rest and at warm-up. We suggest using the smart shirt for monitoring HRV indices on a daily basis, but caution should be taken in the interpretation of HRV indices obtained during moderate to vigorous exercise intensities.

KEYWORDS

bicycle ergometer, comparison study, ECG, heart rate monitor, HRV, incremental exercise test, measurement accuracy

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Clinical Physiology and Functional Imaging* published by John Wiley & Sons Ltd on behalf of Scandinavian Society of Clinical Physiology and Nuclear Medicine.

1 | INTRODUCTION

Heart rate variability (HRV) describes the oscillation of intervals between successive heartbeats (Task Force, 1996). Cardiac functions, and therefore heart rate, are mainly controlled by the autonomic nervous system (Buchheit & Gindre, 2006; Porges, 1992) through modulation in sympathetic and parasympathetic activity (Robinson et al., 1966). The autonomic nervous system in turn is connected to several other physiological systems to maintain the homeostasis in the body by regulating, for example, airway resistance, blood flow, body temperature, energy balance and immune system responses (Wehrwein et al., 2016). Due to these connections, HRV allows to draw useful information about the response of the body to internal and external stressors (Savin et al., 1982) and is sensitive to both physiological and psychological changes (Task Force, 1996). During exercise, an increase in sympathetic activity and a simultaneous reduction in parasympathetic activity occurs (Savin et al., 1982). An opposite shift is observed during recovery after exercise (Perini et al., 1989). Since HRV has the ability to measure the balance between parasympathetic and sympathetic activity (Buchheit & Gindre, 2006), it is suitable for monitoring an athlete's physical stress and recovery.

Measurements of HRV are used to monitor and supervise health and fitness outcomes and have been established in sport and exercise (Singh et al., 2018). HRV provides an objective parameter for coaches and athletes to assess and monitor several aspects: an athlete's fitness level (Djaoui et al., 2017), response to exercise (Guerra et al., 2014), adaptation to training (Plews et al., 2013), recovery status (Chen et al., 2011), overtraining (Le Meur et al., 2013) and response to environmental stressors (Peçanha et al., 2020). Training periodization based on HRV analysis can improve endurance performance effectively (Kiviniemi et al., 2010). HRV can be used to monitor changes in the activity of the autonomic nervous system during exercise (Macartney et al., 2020). In addition, the estimated ventilatory threshold can be determined using HRV measurements and then be applied to plan training load (Grannell & Vito, 2018). Traditionally, recordings of consecutive heartbeats for HRV calculation are performed at rest in a supine position using an electrocardiogram (ECG) detecting time intervals between consecutive R wave peaks (RR intervals) of an ECG trace. Afterwards, HRV is calculated with specialized software. The results are interpreted retrospectively and the cause of changes is assigned to a previous stimulus.

Due to advances in technology, inexpensive and small heart rate monitor devices have been developed that record RR intervals and conform to recommended sampling rate and error handling requirements (Task Force, 1996). In addition, they allow free movement and are easy to handle. Therefore, other applications of interest are becoming feasible, such as data acquisition during exercise like cycling, walking/running, rowing and swimming (McNarry & Lewis, 2012; Michael et al., 2017). Heart rate monitors occur as wireless chest belts, wristbands, rings or watches and are also available as the so-called smart shirts. In general, smart shirts are textiles with embedded electronics or sensors, such as a heart rate monitor. They

stand out for their great comfort and are particularly suitable for everyday use. Smart shirts with heart rate monitors are designed to assess training load during exercises (Romagnoli et al., 2014). They are also used in clinical situations: to improve the diagnosis of cardiovascular diseases in long-term ECG monitoring (Balsam et al., 2018) or during cardiac rehabilitation monitoring in everyday life (Skobel et al., 2014).

At rest, many heart rate monitors and smart shirts display sufficient measurement accuracy reporting HRV by collecting data from ECG signals (Board et al., 2016; Dobbs et al., 2019; Georgiou et al., 2018). However, when subjects were exercising, measurement accuracy was reduced (Dobbs et al., 2019; Georgiou et al., 2018), but may be acceptable for practitioners when monitoring an athlete's HRV (Dobbs et al., 2019). It is assumed that reduced measurement accuracy during exercise is due to increased noise production. Noise production can be caused by the contraction of muscles at the location of sensors, transpiration and movement of the chest during intense breathing (Georgiou et al., 2018). The fit of smart shirts may prevent noise production, allowing for more stable data acquisition (Hong et al., 2009). However, smart shirts' measurement accuracy needs to be evaluated during exercise. To our knowledge, only three studies examined the validity of a smart shirt during exercise for HRV calculation (Heilman & Porges, 2007; Hong et al., 2009; Romagnoli et al., 2014). These three studies showed a promising validity of the examined smart shirts at rest and during exercise, although measurement accuracy declined with increasing exercise intensity (Hong et al., 2009). These studies examined small sample sizes and differed in subject characteristics, exercise type and study protocol. In addition, only one study evaluated the measurement accuracy during different exercise intensities (Hong et al., 2009) and no study during moderate to vigorous exercise intensities. The validity of smart shirts to record RR intervals for HRV calculation during exercise remains questionable.

Therefore, the objective of the present study was to validate the measurement accuracy of a smart shirt for calculating HRV indices at rest and during different exercise intensities. We compared HRV indices calculated from the data recorded by the smart shirt and by ECG in measurement segments at rest and during an incremental exercise test on a bicycle ergometer. We hypothesized that the HRV indices demonstrate (1) a strong linear correlation, (2) an excellent relative reliability and (3) an acceptable agreement.

2 | METHODS

2.1 | Participants

Eighty-three healthy adults aged between 20 and 45 (39 females and 44 males) participated in the present study. Descriptive statistics of the participants are shown in Table 1. All participants were informed of the procedure and the objectives of the study. They signed a consent form in advance, which conformed to the principles of the Declaration of Helsinki for human research of 1974 (last modified in

TABLE 1 Descriptive statistics (mean \pm standard deviation) of the participants

	Women (n = 39)	Men (n = 44)	All (n = 83)
Age (years)	30.9 \pm 6.7	31.3 \pm 6.1	31.0 \pm 6.3
Height (cm)	169.2 \pm 5.2	183.2 \pm 5.9*	176.6 \pm 8.8
Weight (kg)	63.3 \pm 7.8	80.7 \pm 7.2*	72.5 \pm 11.5
Body fat (%)	16.5 \pm 5.1	12.3 \pm 5.2*	14.3 \pm 5.5
Maximal power output (W)	226.4 \pm 38.4	361.4 \pm 57.3*	298 \pm 83.7
Rel. VO_2max ($\text{ml kg}^{-1} \text{min}^{-1}$)	36.9 \pm 6.6	46.1 \pm 6.5*	41.8 \pm 8

Abbreviation: Rel. VO_2max , relative maximal oxygen uptake.

*Statistically significant difference from women ($p < 0.05$) tested with an unpaired t test.

2008). Inclusion and exclusion criteria were collected using a self-report health questionnaire. Inclusion criteria were age between 18 and 45 years, body mass index under 30 and regular participation in sports activities (at least twice a week for 30 min). Exclusion criteria included the use of medication and any metabolic, cardiovascular, musculoskeletal and respiratory disease or illness that could influence the procedure. Participants above 30 years had to present a stress ECG confirmed by a medical doctor that could not be older than 3 months at the time of participation to exclude a cardiac dysfunction. The Ethics Committee of the German Sport University Cologne approved the present study.

2.2 | Experimental protocol

2.2.1 | Setting and instruments

The study procedure was performed in a room with a controlled temperature (20–24°C) and relative humidity (50%–55%). The room was dimly lit during measurements at rest and set to a quiet environment with few distractions. The lights were turned on during the exercise test. The exercise test was performed on the ergoselect 100 bicycle ergometer (ergoline GmbH). The height of the saddle and the handlebar was adapted to the participants. The smart shirt discussed in the present study is named Ambiotex smart shirt (Ambiotex). It was developed for nonprofessional athletes to optimize their daily training and training periodization. It is also used for cardiac long-term monitoring and the detection of mental and physical stress in occupational settings. The examined smart shirt consisted of a heart rate monitor embedded in a tightly fitted but comfortable shirt. Sensors of the embedded heart rate monitor were placed directly under the chest muscles when wearing the shirt. The sensors registered the heart's electrical impulses and an electronic module placed on the chest temporarily saved them. The module transmitted the data in real-time via Bluetooth to an app on an electronic device. The electronic device detected the RR intervals

using an algorithm and then displayed and stored them for later export. Tautinger et al. (2012) tested the algorithm of a previous version of the examined smart shirt in a feasibility study. Ten male participants wore the smart shirt during 2-min segments of sitting, walking, slow jogging, cycling and rowing. Although the sampling rate was only 256 Hz, the smart shirt provided high detection rates, especially during activities with lots of motion.

For comparison to the smart shirt, we used the 5-lead Holter ECG eMotion Faros 360 (Bittium Corporation). Common 10-mm AgCl surface electrodes were placed on the chest according to standard precordial 5-lead positions. ECG's electrodes did not interfere with the smart shirt's sensors. The device recorded the ECG traces and transmitted the data via Bluetooth to the corresponding software on a computer, where the data was analysed later. The ECG device temporarily stored the data in case of connecting problems with the software. Both devices recorded with 1000 Hz sampling rate. Synchronous start of the devices was triggered manually.

2.2.2 | Preparations

Participants were instructed in advance to avoid any strenuous or vigorous exercises 24 h before test, and any alcohol or stimulants (caffeine-related products or with stimulatory effect) 12 h before the test and smoking 2 h before the test. Participants were asked to sleep and drink adequately, to fast 2 h before the test and to wear comfortable clothes. The skin of the participants at the place of device's electrodes was shaved if it was necessary for application.

2.2.3 | Procedure

After acclimatization, the procedure began with a 5-min recording at rest. Therefore, the participants lay in a supine position, moving as little as possible. The incremental exercise test on the bicycle ergometer took place afterwards. Incremental exercise test consisted of two phases. It started with a 3-min warm-up pedalling at a constant load of 10 W, followed by a linear incline of the load until fatigue-related exhaustion. Linear incline of the load was tailored to the participant according to the norm values for the maximal power output of the American College of Sports Medicine (Pescatello et al., 2014). The incline was set for the participants to reach the estimated maximal power output at 10 min of the exercise test. The test protocol aimed at a fatigue-limited exercise duration of 8–12 min to indicate maximal oxygen uptake (VO_2max) (Balady et al., 2010). Cadence was set free between 60 and 80 rpm. Breathing was not controlled and speaking was avoided during measurements.

2.2.4 | Data acquisition and HRV analysis

For data acquisition, smart shirt and ECG simultaneously recorded RR intervals and ECG traces, respectively. Both devices obtained a total

of four segments: at rest (5 min), during warm-up (3 min), the first 5 min of the exercise test starting with the initiation of the increasing load and the last 5 min before exhaustion. The last interval was selected because of individual exercise test durations.

We analysed recorded data of both devices using the HRV-Scanner software (version 3, BioSign). For further HRV analysis, RR intervals were imported from the smart shirt into the software. R wave peak in the traces obtained from the ECG was automatically detected by the software. Both data sets were manually assessed to ensure correct detection of heartbeats according to previous studies (Flatt & Esco, 2015; Weippert et al., 2010). Missed heartbeats in the ECG traces were added manually if the R wave was clearly visible, but not detected. RR intervals from both data sets were excluded if the data had been identified as an artefact, nonsinus origin or within a noisy complex. Consecutive RR intervals greater than 20% were marked as an artefact and removed (Romagnoli et al., 2014). We excluded record segments from further statistical analysis with more than 10% of beat errors (Schäfer & Vagedes, 2013) as well as their paired record segment obtained by the other device.

After data processing, time and frequency-domain HRV indices were calculated. We choose RMSSD, which is defined as the root mean square of successive differences between the RR intervals for time-domain HRV analysis. Fast Fourier transformation was used to calculate frequency-domain HRV indices in different bandwidths. Low-frequency bandwidth was set at 0.04–0.15 Hz and high-frequency bandwidth at 0.15–0.4 Hz (Task Force, 1996). Two frequency-domain HRV indices were calculated in the present study: Relative HF and LF power (rel. HF power, rel. LF power) defined as high-frequency power percentage of total power and low-frequency power percentage of total power, respectively.

2.3 | Statistical analysis

To confirm the validity of the smart shirt, data were analysed in three steps: (1) linear correlation, (2) relative reliability and (3) agreement. Descriptive statistics were calculated for all variables and reported as mean and standard deviation.

(1) For linear correlation analysis, we defined HRV indices calculated from measurement segments obtained by the smart shirt and by ECG as dependent and independent variables, respectively. Each dependent and independent variable was analysed for each of the three HRV indices (RMSSD, rel. HF power and rel. LF power) per one of the four measurement segments separately (rest, warm-up, first 5 min and last 5 min). Pearson's correlation coefficient (r) with 95% confidence intervals (CIs) was used to quantify the linear relationship between the smart shirt and the ECG measurements. In addition, the coefficient of determination (R^2) was calculated to provide the total shared variance between measurements of both devices. The correlation coefficient was set to 0.7, 0.8 and 0.9 as limits for good, high and excellent linear correlation, respectively. A coefficient of determination above 0.7 is considered a strong agreement (Moore et al., 2013). The limit in the present study was set

to 0.9 for an excellent agreement. We hypothesized that Pearson's correlation coefficient and coefficient of determinants demonstrate a strong linear relationship ($r > 0.9$; $R^2 > 0.9$) at rest and during exercise.

(2) Relative reliability was assessed by calculating intraclass correlation coefficients (ICCs) with 95% CIs (Weir, 2005), for which model 3.1 was chosen (Atkinson & Nevill, 1998). Furthermore, Cronbach's α was used for examining internal consistency. We performed both statistical tests for each of the three HRV indices per one of the four measurement segments separately. ICCs higher than 0.8 and 0.9 were considered high and excellent relative reliability, respectively (Weir, 2005). Additionally, ICCs lower 95% CI above 0.8 was defined as tolerable (Pinna et al., 2007). Cronbach's α limit for ideal internal consistency was set above 0.9 (Nunnally & Bernstein, 2008). We hypothesized that ICC with 95% CI and Cronbach's α demonstrate an excellent reliability and internal consistency (ICC > 0.9 ; CI > 0.8 ; $\alpha > 0.9$) at rest and during exercise.

(3) To measure the level of agreement between both devices, we followed the Bland-Altman method comparison. We defined HRV indices calculated from measurement segments obtained by ECG and by the smart shirt as reference variables and as variables to be compared, respectively. Bland-Altman comparison of both variables was performed for each of the three HRV indices per one of the four measurement segments separately. We used the percentage bias and the corresponding lower and upper 95% limits of agreement (Bland & Altman, 2007; Giavarina, 2015). The percentage bias was chosen for better comparison between HRV indices. The limit of acceptable lower and upper percentage bias was set to -10% and 10%, respectively, according to comparable studies (Romagnoli et al., 2014; Weippert et al., 2010). In the Bland-Altman plots, we plotted on the y-axis the percentage differences ($100 \times (\text{ECG} - \text{shirt}) / \text{average}$) against the average values of both ECG and shirt on the x-axis. In addition, mean percentage bias, upper and lower 95% limits of agreement were included in the plots. We hypothesized that Bland-Altman analysis demonstrates an acceptable percentage bias ($10\% < \text{bias} < 10\%$) at rest and during exercise.

Statistical analyses were performed using IBM SPSS Statistics (version 27; IBM Corporation). The level for accepting statistical significance was set at $p < 0.05$ for all analyses. Bland-Altman's comparison and plots were generated with GraphPad Prism (Version 9; GraphPad Software).

3 | RESULTS

Table 2 summarizes the mean, minimum and maximum of the participant's heart rate and power output during the measurement segments. Descriptive statistics of the selected HRV indices measured by ECG and the smart shirt are shown in Table 3, including the mean values and standard deviation. The table also displays the number of record segments for the statistical analysis after excluding segments with excessive error beats. Out of 83 paired records per measurement segment, three paired records were excluded from rest

TABLE 2 Descriptive statistics (mean, minimum, maximum) of heart rate and power output at rest and at warm-up, first 5 min and last 5 min of the exercise test

	Heart rate (bpm)			Power output (W)		
	Mean	Min	Max	Mean	Min	Max
Rest	57	38	75	/	/	/
Warm-up	83	59	121	10.0	10.0	10.0
First 5 min	103	80	140	78.4	49.5	89.0
Last 5 min	155	128	184	227.5	103.5	370.0

Abbreviations: bpm, beats per minute; Min, minimum; Max, maximum; W, Watt.

TABLE 3 Descriptive statistics (mean \pm SD) of HRV indices measured by ECG and smart shirt at rest and at warm-up, first 5 min and last 5 min of the exercise test

	ECG	Smart shirt
Rest ($n = 80$)		
RMSSD (ms)	76.95 \pm 71.64	78.92 \pm 72.68
Rel. HF power (%)	33.94 \pm 20.54	32.77 \pm 20.51
Rel. LF power (%)	36.56 \pm 21.18	38.08 \pm 21.65
Warm-up ($n = 76$)		
RMSSD (ms)	38.10 \pm 30.29	39.77 \pm 30.02
Rel. HF power (%)	23.46 \pm 14.99	22.30 \pm 14.42
Rel. LF power (%)	41.11 \pm 18.41	41.65 \pm 18.76
First 5 min ($n = 76$)		
RMSSD (ms)	17.21 \pm 15.74	19.63 \pm 15.93
Rel. HF power (%)	13.43 \pm 11.55	14.05 \pm 11.11
Rel. LF power (%)	22.01 \pm 12.59	25.37 \pm 14.53
Last 5 min ($n = 78$)		
RMSSD (ms)	3.47 \pm 3.13	4.45 \pm 3.27
Rel. HF power (%)	1.42 \pm 1.87	1.30 \pm 1.96
Rel. LF power (%)	3.48 \pm 6.81	2.61 \pm 6.85

Note: n shows the included number of paired record segments after excluding segments with excessive error beats from a total of 83 possible paired record segments.

Abbreviations: ECG, electrocardiogram; HRV, heart rate variability; Rel. HF power, high-frequency power percentage of total power; rel. LF power, low-frequency power percentage of total power; RMSSD, root mean square of successive differences between the RR intervals; RR, consecutive R wave peaks.

and seven paired records were excluded from warm-up. From the first and last 5 min, seven and five paired records were excluded, respectively.

Table 4 compares the results of (1) linear correlation, (2) relative reliability and (3) agreement analysis. The table includes Pearson's correlation coefficient with 95% CIs and coefficient of determination, ICCs with 95% CIs, Cronbach's α and Bland-Altman's percentage bias

with lower and upper limits of agreement of the calculated HRV indices.

(1) The Pearson's correlation coefficient (r) for the calculated HRV indices with 95% CIs showed an excellent linear relationship ($0.9 < r < 1.1$) between both devices in almost all measurement segments. Correlation coefficient and CIs of rel. LF power fell below the defined limit of 0.9 at the first 5 min of the exercise test. Similarly, the correlation coefficient for RMSSD and rel. HF power declined at the last 5 min of the exercise test. Overall, the linear relationship between both devices declined with increasing load, but remained high ($r > 0.8$). The agreement was excellent ($R^2 > 0.9$) almost overall. Coefficient of determination for RMSSD and rel. HF power declined at the last 5 min of the exercise test, but remained strong ($R^2 > 0.8$). The results of the correlation analysis were all statistically significant ($p < 0.01$).

(2) We observed excellent ICCs (≥ 0.93). Corresponding lower 95% CIs remained high to excellent ($CI > 0.8$). Exceptions were RMSSD at the last 5 min of the exercise test ($CI = 0.77-0.97$). ICCs with CIs were all statistically significant ($p < 0.01$). Cronbach's α showed an excellent consistency ($\alpha \geq 0.94$).

(3) Bland-Altman's limits of agreement showed a sufficiently low percentage bias at rest and at warm-up, and bias ranged from -5.18% to 5.41% . Percentage bias increased at the first 5 min and last 5 min of the exercise test and exceeded the defined limits of percentage bias ($-10\% < \text{bias} < 10\%$). Figures 1-3 show the Bland-Altman plots of RMSSD, rel. HF power and rel. LF power, respectively, at rest and at warm-up, first 5 min and last 5 min of the exercise test. Percentage differences are plotted against the average values of both smart shirt and ECG.

4 | DISCUSSION

We compared the measurement accuracy for HRV calculation of a smart shirt and an ECG during an incremental exercise test on a bicycle ergometer. The findings support the validity of the smart shirt for measuring HRV, although the measurement accuracy declined with increasing exercise intensity. We could confirm our hypotheses at rest and during warm-up of the exercise test: Our results showed (1) an excellent linear relationship, (2) an excellent relative reliability and (3) an acceptable agreement. During the incremental exercise test, the results for partial HRV indices fell under the defined limits but remained high. However, percentage bias was unacceptable during the exercise test.

At rest and during warm-up of the exercise test, the measurement accuracy of the examined smart shirt for calculating HRV indices was sufficient. We found an excellent linear relationship ($r > 0.96$; $R^2 > 0.94$), an excellent relative reliability ($ICC \geq 0.98$; lower $CI \geq 0.97$; $\alpha \geq 0.98$) and an acceptable agreement (bias $< 10\%$). These findings are comparable with the results of two studies, which examined the measurement accuracy of smart shirts during similar intensities. Romagnoli et al. (2014) contrasted a smart shirt and an ECG during a 30-min bicycle ergometer exercise at a stable

TABLE 4 Comparison of Pearson's correlation coefficient (r) with 95% CIs and coefficient of determination (R^2), ICCs with 95% CIs, Cronbach's α and percentage bias with lower and upper LoA of heart rate variability indices at rest and at warm-up, first 5 min and last 5 min of the exercise test

	Pearson's correlation		Reliability		Bland-Altman analysis Percentage bias (LoA)
	r (CI)	R^2	ICC (CI)	α	
Rest					
RMSSD (ms)	0.980 (0.957–1.004)	0.989	1.00 (1.00–1.00)	1.00	2.79 (–16.00 to 21.57)
Rel. HF power (%)	0.990 (0.956–1.025)	0.977	0.99 (0.99–1.00)	0.99	–5.18 (–30.28 to 19.91)
Rel. LF power (%)	0.972 (0.947–0.997)	0.987	1.00 (0.99–1.00)	1.00	4.78 (–13.11 to 22.67)
Warm-up					
RMSSD (ms)	1.002 (0.976–1.029)	0.987	1.00 (0.99–1.00)	1.00	5.41 (–9.02 to 19.83)
Rel. HF power (%)	1.008 (0.949–1.067)	0.941	0.98 (0.97–0.99)	0.98	–4.04 (–39.60 to 31.52)
Rel. LF power (%)	0.962 (0.918–1.007)	0.961	0.99 (0.98–0.99)	0.99	0.73 (–29.36 to 30.81)
First 5 min					
RMSSD (ms)	0.962 (0.909–1.014)	0.947	0.98 (0.94–0.99)	0.99	15.80 (–22.99 to 54.59)
Rel. HF power (%)	1.016 (0.964–1.067)	0.954	0.99 (0.98–0.99)	0.99	10.02 (–36.13 to 56.18)
Rel. LF power (%)	0.826 (0.765–0.887)	0.908	0.96 (0.84–0.98)	0.97	14.23 (–30.98 to 59.45)
Last 5 min					
RMSSD (ms)	0.870 (0.780–0.961)	0.829	0.93 (0.77–0.97)	0.95	26.84 (–35.51 to 89.18)
Rel. HF power (%)	0.856 (0.760–0.953)	0.804	0.95 (0.91–0.96)	0.94	–7.62 (–103.50 to 88.24)
Rel. LF power (%)	0.969 (0.918–0.020)	0.950	0.98 (0.96–0.99)	0.99	–27.82 (–95.31 to 39.66)

Note: Results of r (CI), R^2 and ICC (CI) were all statistically significant ($p < 0.01$).

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficients; LoA, limits of agreement; rel. HF power, high-frequency power percentage of total power; rel. LF power, low-frequency power percentage of total power; RMSSD, root mean square of successive differences between the RR intervals.

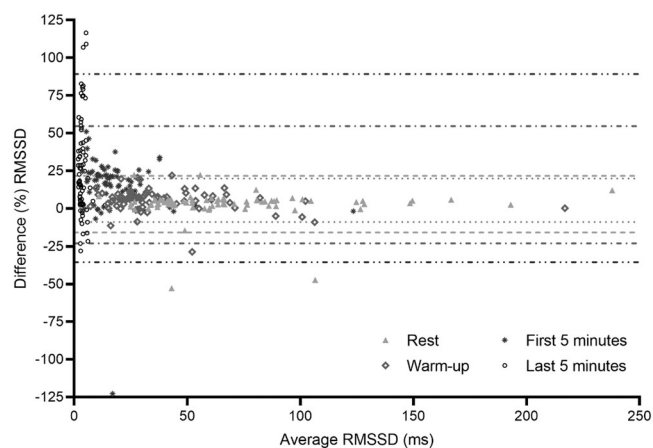


FIGURE 1 Bland-Altman plots for RMSSD at rest and at warm-up, first 5 min and last 5 min of the exercise test. Percentage differences ($100 \times (\text{ECG} - \text{shirt})/\text{average}$) are plotted against the average values of both ECG and smart shirt. Dashed lines show mean percentage bias and dotted lines show upper and lower 95% limits of agreement. X-axis was adjusted to fit the data for better display. ECG, electrocardiogram; RMSSD, root mean square of successive differences between the RR intervals; RR, consecutive R wave peaks

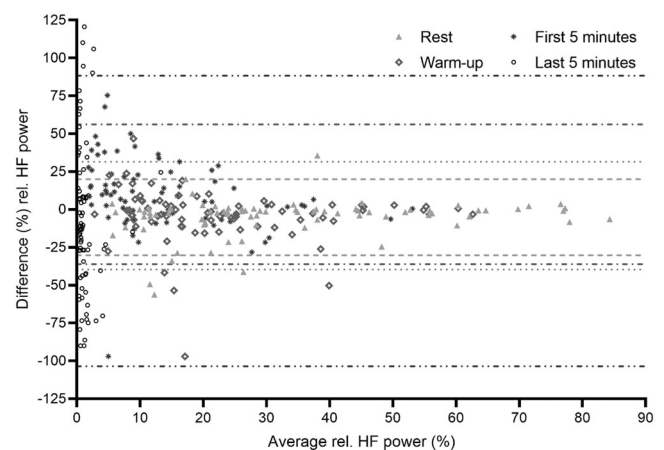


FIGURE 2 Bland-Altman plots for rel. HF power at rest and at warm-up, first 5 min and last 5 min of the exercise test. Percentage differences ($100 \times (\text{ECG} - \text{shirt})/\text{average}$) are plotted against the average values of both ECG and smart shirt. Dashed lines show mean percentage bias and dotted lines show upper and lower 95% limits of agreement. X-axis was adjusted to fit the data for better display. ECG, electrocardiogram; rel. HF power, high-frequency power percentage of total power

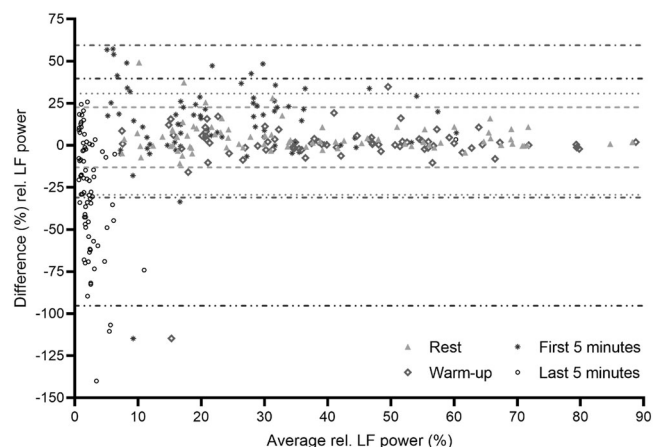


FIGURE 3 Bland–Altman plots for rel. LF power at rest and at warm-up, first 5 min and last 5 min of the exercise test. Percentage differences ($100 \times (\text{ECG} - \text{shirt}) / \text{average}$) are plotted against the average values of both ECG and smart shirt. Dashed lines show mean percentage bias and dotted lines show upper and lower 95% limits of agreement. X-axis was adjusted to fit the data for better display. ECG, electrocardiogram; rel. LF power, low-frequency power percentage of total power

submaximal intensity (HR-leaded at 75–95 bpm). The authors calculated HRV indices RMSSD, LF power and HF power from ten 3-min segments. Results showed an ICC and lower 95% CI above 0.9. Bland–Altman's percentage bias and limits of agreement were acceptable (<10%). However, RMSSD and HF power partial failed the defined limits of the authors. Although the study protocol was similar, the measurement accuracy of the examined smart shirt in the present study seems more accurate. Hong et al. (2009) compared HRV indices HF power and LF power obtained by a smart shirt and an ECG during walking or running for 5 min on a treadmill at variable speeds. During 3 and 6 km/h both devices showed a linear correlation coefficient and lower 95% CI above 0.9. Although the exercise differed compared to the present study, the findings of linear correlation analysis were similar. Furthermore, the smart shirt showed a similar validity like chest belts at rest (Da Costa de Rezende Barbosa et al., 2014) and during light exercise on a bicycle ergometer (Kingsley et al., 2005). Therefore, we assume that athletes can use the smart shirt to determine their recovery status at rest and monitor their training load during light exercise.

Regarding the first 5 min of the exercise test in the present study, linear correlation and relative reliability remained excellent except for rel. LF power. However, agreement analysis showed unacceptable limits of agreements for all calculated HRV indices ($-10\% > \text{bias} > 10\%$). A similar decrease in correlation coefficient for frequency-domain HRV indices between a smart shirt and an ECG was shown during running at a speed of 9 km/h (Hong et al., 2009). To our knowledge, there was no other study examining the measurement accuracy for HRV calculation of a smart shirt during similar exercise intensities. Comparing a chest belt and an ECG during exercise, two studies used a comparable graded exercise test on a bicycle

ergometer (Hernando et al., 2018; Kingsley et al., 2005). The present study and both mentioned studies showed similar results. When the participants exercised at intensities up to 80% of VO_2max , linear correlation and ICC were sufficient for the HRV index LF power, but declined for HF power (Hernando et al., 2018) and Bland–Altman's percentage limits of agreement became unacceptable for both HRV indices (Kingsley et al., 2005). Both author groups assumed that caution should be taken when interpreting frequency-domain HRV indices during exercise. Due to the unacceptable percentage bias between both examined devices in the present study, this assumption can be confirmed.

It should be noted that measurement accuracy of heart rate monitors depends on an accurate algorithm for recording and detecting the R wave peak within the ECG traces (Kingsley et al., 2005; Weippert et al., 2010) and on artefacts correction (Akintola et al., 2016). Single artefacts and even single artificially corrected beats are known to have effects on short-term HRV indices and especially on frequency-domain HRV indices (Buchheit, 2014; Clifford & Tarassenko, 2005). The differences between the results of the calculated HRV indices by Kingsley et al. (2005), Hernando et al. (2018) and the present study may be caused by different data acquisitions. Similarly, within the present study, differences between HRV indices could have occurred due to varied data acquisition. The smart shirt automatically detected the R waves and only RR intervals could be exported for calculating HRV indices, whereas the ECG recorded the traces of the heart's electrical actions and R wave detection could be reviewed visually before HRV calculation. Therefore, manual control of the collected data should be preferred over processing obtained RR intervals for accurate HRV calculation. However, the level of agreement between the smart shirt and ECG during exercise may be acceptable for practitioners when monitoring HRV. Since the examined smart shirt was developed for nonprofessional athletes, the benefit of automatic data processing might outweigh the limited measurement accuracy at this level of proficiency.

When participants exercised at the last 5 min of the exercise test, linear correlation of RMSSD and rel. HF power fell below the defined limit. ICC's lower 95% confidence interval for RMSSD also fell below the defined limit. In addition, Bland–Altman's percentage lower and upper limits of agreement were unacceptable ranging from -103.5% to 89.18% respectively. Similar results were found by Hernando et al. (2018) while comparing a chest belt with an ECG for recording HRV data during a graded exercise test on a bicycle ergometer at intensities above 60% of VO_2max . Correlation coefficient and ICC with 95% CI for HF power were insufficient. Results for LF power remained sufficient at this level of exercise. The same trend can be seen in the present study for rel. HF power and rel. LF power at the last 5 min during the exercise test, although results of ICC for rel. HF power remained high and therefore seemed to be more stable than the results of Hernando et al. (2018). Kingsley et al. (2005) found in their examination of measurement accuracy of a chest belt during a graded exercise test on a bicycle ergometer a broader range of the percentage limits of agreement. Lower and higher limits of agreement for frequency-domain HRV indices HF and LF power ranged from -1481.8% to 388.7% at intensities above 60% of VO_2max . A broader

range of the percentage limits of agreement may have occurred due to the use of absolute values for HF and LF power in contrast to the use of relative values in the present study. Therefore, supposedly stable indices should be used for HRV monitoring like relative or logarithmic indices. RMSSD and HF power estimate short-term components of variation in RR intervals (Task Force, 1996). Therefore, both indices are sensitive for changes in RR interval length caused by motion artefacts, ectopic beats or data correction.

Overall, we observed the trend that measurement accuracy of the smart shirt for HRV monitoring declined progressively as exercise intensity increased during the incremental exercise test. The same trend was found in validation studies of heart rate monitors during graded exercise test on bicycle ergometers. Kingsley et al. (2005) observed no differences between a chest belt and an ECG at rest and during low exercise intensity. However, when the intensity increased further, a progressive decline was observed, especially for frequency-domain HRV indices at an exercise intensity above 60% of VO_2max . Hernando et al. (2018) found a decline in measurement accuracy for frequency-domain HRV indices when RR intervals were shorter. Also while walking and running at different speeds, a progressive decline of a smart shirt's measurement accuracy was observed by Hong et al. (2009). Their results suggested a more stable accuracy by smart shirt than by ECG caused by less noise production. It is assumed that declined measurement accuracy occurs due to noise production by movement at the sensor placement (Georgiou et al., 2018). Therefore, care should be taken during RR interval measurement to minimize unnecessary movement at the sensor placement while exercising.

The present study has three main limitations. First, participants performed the incremental exercise test on a bicycle ergometer. By bracing the arms on the handlebars of the ergometer during the exercise test, the upper body was probably additionally stabilized. Due to the stabilized upper body, we assumed that less motion of the sensor and less noise production occurred than it could be the case while running. Therefore, the practical advice is limited to exercises with less upper body movements. Second, the synchronization of the RR interval recording was manually set for both devices. The possible delay may have caused differences in the starting time and therefore included different RR intervals for HRV calculation. However, since the results were all statistically significant, the differences seem to be negligible. Third, noise production could be caused by the motion of the smart shirts' sensors and the ECG electrodes separately. Due to the different placements of the sensors and the electrodes, varied noise productions could have occurred and led to different RR interval recordings. Therefore, an interpretation of which device provides more stable recordings is difficult to achieve.

5 | CONCLUSION

To our knowledge, this is the first study that compared a smart shirt and an ECG in an incremental exercise test for HRV calculations, especially at moderate to vigorous exercise intensities. The results promote that both devices are interchangeable when analysing HRV

at rest and during light exercise. The agreement between both devices declined with increasing intensity, but remained high. The examined smart shirt was developed for nonprofessional athletes. The advantages of this smart shirt include the easy use and comfortable application compared to an ECG. Therefore, the validity of this smart shirt may be sufficient for the use of nonprofessional athletes. Since alternative devices such as chest belts show a similar validity, the user has to decide if smart shirts' advantages outweigh the price difference. In conclusion, the smart shirt is a valid device for monitoring HRV while performing physical exercise on a daily basis, although caution should be taken in the interpretation of HRV indices obtained during moderate to vigorous exercise intensities.

ACKNOWLEDGEMENTS

The present study is fully funded by the Central Innovation Programme for small and medium-sized enterprises of the Federal Ministry for Economic Affairs and Energy of Germany (funding code: ZF 4423401AK7). We would like to thank the athletes for volunteering their time to participate in the present study. Open access funding enabled and organized by Projekt DEAL.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Raw data were generated at the Institute of Movement Therapy and Movement-oriented Prevention and Rehabilitation Sciences of the German Sport University Cologne. The data that support the findings of this study are available from the corresponding author upon reasonable request. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Peter Schams  <http://orcid.org/0000-0001-9095-3195>

Boris Feodoroff  <http://orcid.org/0000-0002-5169-9005>

REFERENCES

- Akintola, A.A., van de Pol, V., Bimmel, D., Maan, A.C. & van Heemst, D. (2016) Comparative analysis of the Equivital EQ. 02 Lifemonitor with Holter ambulatory ECG device for continuous measurement of ECG, heart rate, and heart rate variability: a validation study for precision and accuracy. *Frontiers in Physiology*, 7, e1–e14.
- Atkinson, G. & Nevill, A.M. (1998) Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*, 26, 217–238.
- Balady, G.J., Arena, R., Sietsema, K., Myers, J., Coke, L., Fletcher, G.F. et al. (2010) Clinician's guide to cardiopulmonary exercise testing in adults: a scientific statement from the American Heart Association. *Circulation*, 122, 191–225.
- Balsam, P., Lodziński, P., Tymińska, A., Ozierański, K., Januszkiwicz, Ł., Główczyńska, R. et al. (2018) Study design and rationale for biomedical shirt-based electrocardiography monitoring in relevant clinical situations: ECG-shirt study. *Cardiol J*, 25, 52–59.
- Bland, J.M. & Altman, D.G. (2007) Agreement between methods of measurement with multiple observations per individual. *Journal of Biopharmaceutical Statistics*, 17, 571–582.

- Board, E.M., Ispoglou, T. & Ingle, L. (2016) Validity of telemetric-derived measures of heart rate variability: a systematic review. *Journal of Exercise Physiology Online*, 19, 64–84.
- Buchheit, M. (2014) Monitoring training status with HR measures: do all roads lead to Rome? *Frontiers in Physiology*, 5, e1–e19.
- Buchheit, M. & Gindre, C. (2006) Cardiac parasympathetic regulation: respective associations with cardiorespiratory fitness and training load. *American Journal of Physiology: Heart and Circulatory Physiology*, 291, H451–H458.
- Chen, J.-L., Yeh, D.-P., Lee, J.-P., Chen, C.-Y., Huang, C.-Y., Lee, S.-D. et al. (2011) Parasympathetic nervous activity mirrors recovery status in weightlifting performance after training. *Journal of Strength and Conditioning Research*, 25, 1546–1552.
- Clifford, G.D. & Tarassenko, L. (2005) Quantifying errors in spectral estimates of HRV due to beat replacement and resampling. *IEEE Transactions on Biomedical Engineering*, 52, 630–638.
- Da Costa de Rezende Barbosa, M.P., da Silva, N.T., de Azevedo, F.M., Pastre, C.M. & Vanderlei, L. (2014) Comparison of Polar®RS800G3™ heart rate monitor with Polar®S810i™ and electrocardiogram to obtain the series of RR intervals and analysis of heart rate variability at rest. *Clinical Physiology and Functional Imaging*, 36, 112–117.
- Djaoui, L., Haddad, M., Chamari, K. & Dellal, A. (2017) Monitoring training load and fatigue in soccer players with physiological markers. *Physiology and Behavior*, 181, 86–94.
- Dobbs, W.C., Fedewa, M.V., MacDonald, H.V., Holmes, C.J., Cicone, Z.S., Plews, D.J. et al. (2019) The accuracy of acquiring heart rate variability from portable devices: a systematic review and meta-analysis. *Sports Medicine*, 49, 417–435.
- Flatt, A.A. & Esco, M.R. (2015) Heart rate variability stabilization in athletes: towards more convenient data acquisition. *Clinical physiology and functional imaging*, 36, 331–336.
- Georgiou, K., Larentzakis, A.V., Khamis, N.N., Alsuhaibani, G.I., Alaska, Y.A. & Giallafos, E.J. (2018) Can wearable devices accurately measure heart rate variability? a systematic review. *Folia Medica (Plovdiv)*, 60, 7–10.
- Giavarina, D. (2015) Understanding Bland–Altman analysis. *Biochem Med (Zagreb)*, 25, 141–151.
- Grannell, A. & De Vito, G. (2018) An investigation into the relationship between heart rate variability and the ventilatory threshold in healthy moderately trained males. *Clinical Physiology and Functional Imaging*, 38, 455–461.
- Guerra, Z.F., Peçanha, T., Moreira, D.N., Silva, L.P., Laterza, M.C., Nakamura, F.Y. et al. (2014) Effects of load and type of physical training on resting and postexercise cardiac autonomic control. *Clinical Physiology and Functional Imaging*, 34, 114–120.
- Heilman, K.J. & Porges, S.W. (2007) Accuracy of the LifeShirt (Vivometrics) in the detection of cardiac rhythms. *Biological Psychology*, 75, 300–305.
- Hernando, D., Garatachea, N., Almeida, R., Casajús, J.A. & Bailón, R. (2018) Validation of heart rate monitor Polar RS800 for heart rate variability analysis during exercise. *Journal of Strength and Conditioning Research*, 32, 716–725.
- Hong, S., Yang, Y., Kim, S., Shin, S., Lee, I., Jang, Y. et al. (2009) Performance study of the wearable one-lead wireless electrocardiographic monitoring system. *Telemedicine Journal and e-Health: The Official Journal of the American Telemedicine Association*, 15, 166–175.
- Kingsley, M., Lewis, M.J. & Marson, R.E. (2005) Comparison of Polar 810 s and an ambulatory ECG system for RR interval measurement during progressive exercise. *International Journal of Sports Medicine*, 26, 39–44.
- Kiviniemi, A.M., Hautala, A.J., Kinnunen, H., Nissilä, J., Virtanen, P., Karjalainen, J. et al. (2010) Daily exercise prescription on the basis of HR variability among men and women. *Medicine and Science in Sports and Exercise*, 42, 1355–1363.
- Le Meur, Y., Pichon, A., Schaal, K., Schmitt, L., Louis, J., Gueneron, J. et al. (2013) Evidence of parasympathetic hyperactivity in functionally overreached athletes. *Medicine and Science in Sports and Exercise*, 45, 2061–2071.
- Macartney, M.J., Meade, R.D., Notley, S.R., Herry, C.L., Seely, A.J.E. & Kenny, G.P. (2020) Fluid loss during exercise-heat stress reduces cardiac vagal autonomic modulation. *Medicine and Science in Sports and Exercise*, 52, 362–369.
- McNarry, M.A. & Lewis, M.J. (2012) Heart rate variability reproducibility during exercise. *Physiological measurement*, 33, 1123–1133.
- Michael, S., Graham, K.S. & Davis, G.M. (2017) Cardiac autonomic responses during exercise and post-exercise recovery using heart rate variability and systolic time intervals—a review. *Frontiers in Physiology*, 8, e1–e19.
- Moore, D.S., Notz, W. & Fligner, M.A. (2013) *The basic practice of statistics*, 6th edition, New York: W.H. Freeman and Co.
- Nunnally, J.C. & Bernstein, I.H. (Eds.) (2008) *Psychometric theory*, 3rd edition, New York: McGraw-Hill.
- Peçanha, T., Low, D., de Brito, L.C., Fecchio, R.Y., de Sousa, P.N. & da Silva-Júnior, N.D. et al. (2020) Effects of postexercise cooling on heart rate recovery in normotensive and hypertensive men. *Clinical Physiology and Functional Imaging*, 40, 114–121.
- Perini, R., Orizio, C., Comandè, A., Castellano, M., Beschi, M. & Veicsteinas, A. (1989) Plasma norepinephrine and heart rate dynamics during recovery from submaximal exercise in man. *European Journal of Applied Physiology and Occupational Physiology*, 58, 879–883.
- Pescatello, L.S., Arena, R., Riebe, D. & Thompson, P.D. (2014) *ACSM's guidelines for exercise testing and prescription*, 9th edition, Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins.
- Pinna, G.D., Maestri, R., Torunski, A., Danilowicz-Szymanowicz, L., Szwoch, M., La Rovere, M.T. et al. (2007) Heart rate variability measures: a fresh look at reliability. *Clinical Science (London)*, 113, 131–140.
- Plews, D.J., Laursen, P.B., Stanley, J., Kilding, A.E. & Buchheit, M. (2013) Training adaptation and heart rate variability in elite endurance athletes: opening the door to effective monitoring. *Sports Medicine*, 43, 773–781.
- Porges, S.W. (1992) Vagal tone: a physiologic marker of stress vulnerability. *Pediatrics*, 90, 498–504.
- Robinson, B.F., Epstein, S.E., Beiser, G.D. & Braunwald, E. (1966) Control of heart rate by the autonomic nervous system. Studies in man on the interrelation between baroreceptor mechanisms and exercise. *Circulation Research*, 19, 400–411.
- Romagnoli, M., Alis, R., Guillen, J., Basterra, J., Villacastin, J.P. & Guillen, S. (2014) A novel device based on smart textile to control heart's activity during exercise. *Australasian Physical and Engineering Sciences in Medicine*, 37, 377–384.
- Savin, W.M., Davidson, D.M. & Haskell, W.L. (1982) Autonomic contribution to heart rate recovery from exercise in humans. *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology*, 53, 1572–1575.
- Schäfer, A. & Vagedes, J. (2013) How accurate is pulse rate variability as an estimate of heart rate variability? a review on studies comparing photoplethysmographic technology with an electrocardiogram. *International Journal of Cardiology*, 166, 15–29.
- Singh, N., Moneghetti, K.J., Christle, J.W., Hadley, D., Froelicher, V. & Plews, D. (2018) Heart rate variability: an old metric with new meaning in the era of using mhealth technologies for health and exercise training guidance. Part two: prognosis and training. *Arrhythmia & Electrophysiology Review*, 7, 247–255.
- Skobel, E., Martínez-Romero, A., Scheibe, B., Schauer, P., Marx, N., Luprano, J. et al. (2014) Evaluation of a newly designed shirt-based

- ECG and breathing sensor for home-based training as part of cardiac rehabilitation for coronary artery disease. *European Journal of Preventive Cardiology*, 21, 1332–1340.
- Tantinger, D., Feilner, S., Schmitz, D., Weigand, C., Hofmann, C. & Struck, M. (2012) Evaluation of QRS detection algorithm implemented for mobile applications based on ECG data acquired from sensorized garments. *Biomedizinische Technik*, 57, 635–638.
- Task Force of the European Society of Cardiology & the North American Society of Pacing and Electrophysiology. (1996) Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation*, 93, 1043–1065.
- Wehrwein, E.A., Orer, H.S. & Barman, S.M. (2016) Overview of the anatomy, physiology, and pharmacology of the autonomic nervous system. *Comprehensive Physiology*, 6, 1239–1278.
- Weippert, M., Kumar, M., Kreuzfeld, S., Arndt, D., Rieger, A. & Stoll, R. (2010) Comparison of three mobile devices for measuring R–R intervals and heart rate variability: Polar S810i, Suunto t6 and an ambulatory ECG system. *European Journal of Applied Physiology and Occupational Physiology*, 109, 779–786.
- Weir, J.P. (2005) Quantifying test–retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research*, 19, 231–240.

How to cite this article: Schams, P., Feodoroff, B., Zacher, J., Eibl, A. & Froböse, I. (2022) Validation of a smart shirt for heart rate variability measurements at rest and during exercise. *Clinical Physiology and Functional Imaging*, 42, 190–199. <https://doi.org/10.1111/cpf.12746>