Journal of Sports Sciences

ISSN: 0264-0414 (Print) 1466-447X (Online) Journal homepage: http://www.tandfonline.com/loi/rjsp20

Accuracy of the wearable activity tracker Garmin Forerunner 235 for the assessment of heart rate during rest and activity

Morten Pallisgaard Støve, Emil Haucke, Maria Lysbeck Nymann, Tordur Sigurdsson & Birgit Tine Larsen

JOURNAL OF SPORTS SCIENCES
https://doi.org/10.1080/02640414.2018.1535563

PHYSICAL ACTIVITY, HEALTH AND EXERCISE

Accuracy of the wearable activity tracker Garmin Forerunner 235 for the assessment of heart rate during rest and activity

Morten Pallisgaard Støve, Emil Haucke, Maria Lysbeck Nymann, Tordur Sigurdsson and Birgit Tine Larsen

Department of Physiotherapy, University College of Northern Denmark (UCN), Aalborg Ø, Denmark

ABSTRACT

Accurate measures of heart rate (HR) during rehabilitation and sporting activities are important for precise exercise prescription to maintain or increase capacity. Wrist-worn activity monitors utilizing photoplethysmography technology (PPG) to configure HR show discrepant findings regarding validity depending on the type and intensity of exercises measured, and no previous study has yet investigated the accuracy during running at speeds exceeding 9.6 km/h. The purpose of the study was to assess the accuracy of the Garmin Forerunner 235 (GF), at different exercises at various intensities. Twenty-nine participants participated in the study. HR was measured with the (GF) during rest and three submaximal exercise conditions; cycling, treadmill walking, running and rapid arm movements. The GF had high agreement with the PL during rest (r = 0.997) cycling at 150 W (Rho = 0.889), treadmill running at 8.7 km/h (r = 0.906) and 12.1 km/h (r = 0.845) and rapid arm movements (r = 0.928, r = 0.745) but a low agreement during cycling at 50 W (Rho = 0.269) and 100W (Rho = 0.462) and treadmill walking at 4.8 km/h (r = 0.481). The results varied across exercise conditions and intensities and although the GF provided accurate measurements of HR during rest, cycling at 150W, treadmill running, and rapid arm movement measurement latency may potentially affect application.

ARTICLE HISTORY
Accepted 8 October 2018

KEYWORDS
Activity monitors; validity; exercise; heart rate monitoring

Introduction

Heart rate (HR) monitoring is a useful tool for determining, assessing and evaluating exercise intensity and individual cardiovascular activity (Engström, Ottosson, Wohlfart, Grundström, & Wisén, 2012; Weippert et al., 2010) and is widely used in recreational, experimental and clinical settings (Vasconcellos et al., 2015).

HR monitoring is specifically used by physiotherapists to assess individual capacity and exercise intensity as well as to motivate patients to exercise at specific intensities in order to maintain or increase capacity (Engström et al., 2012), hence accurate measures of HR are important for precise exercise prescription (Wallen, Gomersall, Keating, Wisløff, & Coombes, 2016).

Several technologies such as ECG and cardiotachometers (e.g., pulse rate monitors utilizing a chest strap with an in-built transmitter) have previously been used to assess HR (Vasconcellos et al., 2015) and traditional pulse rate monitors utilizing chest straps show excellent validity and reliability in both rest and exercise conditions (Gillinov et al., 2017; Parak & Korhonen, 2014). They can, however, be inconvenient to use, as the chest strap needs to be worn on the skin and kept wet for accurate signal detection, hence they may cause minor discomfort when worn for extended periods (Fallow, Tarumi, & Tanaka, 2013; Spierer, Rossen, Litman, & Fuji, 2015).

Advances in technology have led to the emergence of new consumer-based wrist-worn activity monitors that utilizes photoplethysmography technology (PPG) to measure HR. PPG sensors use LED emitted light to detect changes in pulsatile blood flow between systole and diastole to configure HR (Fallow et al., 2013). Questions regarding the accuracy of wearable PPG devices are increasingly relevant considering the current shift in consumer transition from the traditional chest strap based monitors to wearable PPG devices (Gillinov et al., 2017). Although wrist-worn activity monitors demonstrate good potential (Lee, Kim, & Welk, 2014) studies show somewhat discrepant findings regarding its validity (Lee et al., 2014; Stahl, An, Dinkel, Noble, & Lee, 2016; Wallen et al., 2016) depending on the devices used and the type and intensity of activity measured (Gillinov et al., 2017; Wallen et al., 2016). Some studies suggest that wrist-worn activity monitors perform best at low exercise intensities while others suggest good accuracy even at near-maximal exercise intensities (Alzahrani et al., 2015; Stahl et al., 2016). Movement of the hand and forearm has been linked with reduced PPG sensor accuracy and current evidence suggests that wearable PPG sensors may perform best during exercise activities with little or no hand/forearm movement (Gillinov et al., 2017; Spierer et al., 2015). Although several studies have investigated the accuracy of wrist-worn activity monitors utilizing PPG technology to configure HR at running speeds up to 9.6 km/h (Gillinov et al., 2017; Spierer et al., 2015), no single study has assessed the accuracy of wrist-worn activity monitors utilizing PPG technology at specific running speeds exceeding 9.6 km/h even though most recreational runners report mean running speeds exceeding 10 km/h (Forsberg, Analyseinstitut, Idræts-Forbund, & Forbund, 2012; Malisoux, Nielsen, Urhausen, & Theisen, 2015). Little is thus known about the accuracy of wearable PPG technology during running at higher intensities.

CONTACT Morten Pallisgaard Støve mps@ucn.dk University College of Northern Denmark, Selma Lagerløfs vej 2, 9220 Aalborg Ø, Denmark

The purpose of this study was to investigate the concurrent validity of HR measured by the Garmin Forerunner 235, a wrist-worn activity monitor at different activities at varying intensities including running at speeds exceeding 9.6 km/h. The overall aim was to determine whether the Garmin Forerunner 235 is suitable for measuring HR during physical activity and exercise training.

Methods

In this cross-sectional study HR was systematically assessed in a single session for each participant with the purpose of assessing the accuracy of the Garmin Forerunner 235, compared with the criterion measure Polar RS400, during rest and three different submaximal to near-maximal exercise intensity level tests with and without hand and forearm movements (e.g., stationary cycling, treadmill walking/running and rapid arm movements).

Participants

Twenty-nine healthy participants were recruited for this study. The sample size was determined based upon findings from similar studies (de Rezende Barbosa, Silva, de Azevedo, Pastre, & Vanderlei, 2016; Nunan et al., 2009; Wallen et al., 2016). Participants were eligible if they were between the ages of 18 and 60 years. Exclusion criteria were a history of injury or disease that would prevent participants from safely performing the study protocol, the use of medications influencing metabolic, endocrine or cardiovascular systems or being a smoker. All participants were asked to refrain from alcohol and physical exercises 24 hours prior to participation (Nunan et al., 2009) and abstain from food intake and beverages that contain caffeine, for 2 h prior to participation (Engström et al., 2012). The participants entered the study after providing their written informed consent. The study was conducted in accordance with the declaration of Helsinki however; ethical approval was not needed according to The North Denmark Regional Committee on Health Research Ethics (02.11.2016). All testing was completed at the musculoskeletal laboratory at the University College of Northern Denmark, with four unblinded investigators performing all tests.

Procedures

Participants performed a standardized exercise protocol in a controlled laboratory setting. The activities and intensities were selected to represent submaximal to near-maximal exercise intensities with and without hand/forearm movement. The activities, presented in experimental order, with a standardized transition time of 3 minutes between activities, were as follows: 1) Rest; participants placed supine on a bed, eyes closed, with the head supported by a pillow, for 10 minutes (Nunan et al., 2009). 2) Participants transitioned to a cycle ergometer (Monark 828E, Varberg, Sweden), adjusted according to individual height specifications. Prior to measurements, each participant was instructed to cycle with a pedal rate of 50 rpm at a load of 25 W for a 2-minute warm-up period (Engström et al., 2012). Participants then cycled for a period of three minutes at each of three power levels: 50, 100 and 150 W, with no rest in between power levels. 3) Participants transitioned to a treadmill where they were instructed to walk at 4.8 km/h and run at 8.7 and 12.1 km/h for 3 minutes respectively. Upon completion, participants cooled down by walking at 4.8 km/h for 2 minutes. 4) This lead to a final transition to a rapid arm movement protocol where participants carried out light weight lifting (1 kg for women and 2 kg for men) comprising of moving the arms to a biceps curl position then to a bilateral "press" over the head and then returning to the start position. Participants completed one cycle of weight lifting per beat to the sound of a metronome set at 40 beats per minute for one minute. (Spierer et al., 2015) Participants completed two sets of weight lifting with a one-minute resting period in between sets. Simultaneous HR measurements were made every minute and data from the last measurement in each activity level was used for analysis.

Instrumentation

The Polar RS400 (Polar Electro Oy, Kempele, Finland) (PL) was used as criterion measure and utilizes a chest strap, with an inbuilt transmitter, that detects the QRS-complexes with 1-ms resolution and sends an electromagnetic signal to a wrist-worn watch that measures the R–R interval which form the basis for the calculation of HR in beats per minute (bpm). The Polar RS400 has been found valid and reliable during both rest and exercise. (Engström et al., 2012)

HR was measured using the Garmin Forerunner 235 (Garmin International Inc, Olathe, KS, USA.) (GF) a wrist-worn activity tracker utilizing green LED optical sensors to measure the amount of light refracted in the blood vessels utilizing the PPG technique to calculate the HR in bpm. (Stahl et al., 2016) Prior to applying the monitors participants were cleaned with a disinfectant wipe around the thorax and on the left wrist to remove any topical creams or lotions (Spierer et al., 2015). Participants were fitted with the wrist-worn GF placed on the left wrist according to manual specifications. The PL chest strap was wetted and fitted around the thorax according to manual specifications and the wrist receiver was placed within 1 ft of the participant so that HR could be manually recorded (Stahl et al., 2016) (Figure 1). As per manufacturer instructions, the devices were individualized for age, gender, and anthropometrical data prior to the start of the measurements (Wallen et al., 2016) as the same chest strap and wrist units were utilized by all test participants. HR was concurrently assessed with both monitors and manually recorded by a researcher taking a digital picture every 60 seconds with both the GF and PL monitors' in the same frame thus ensuring that criterion measures were obtained simultaneously as per following procedure. During cycling, the PL was mounted on the handlebar, as shown in Figure 1 (Figure 1). During walking and running the PL was mounted on the treadmill handlebar. Participants were verbally instructed to briefly place their left hand on the handlebar, next to the PL, while walking or running. During rapid arm movement, the participants were instructed to briefly extend their arm, where a researcher, holding the PL, would

Figure 1. Subject outfitted with the Garmin Forerunner 235 on the left wrist and criterion measure Polar RS400 monitor placed at the cycle ergometer handle for simultaneous HRV measures.

place it next to the GF. Following confirmation that both GF and PL heart rate indicators were readable in the digital picture, the participant would then be instructed to remove their hand from the handlebar or retract their arm. This process lasted approximately 5 seconds per measurement.

Statistical analyses

The data were analysed using SPSS 23 (*SPSS Inc., Chicago, IL, USA*). Normality of the data was assessed by the Shapiro-Wilks test. The minimal detectable change (MDC_{95} = 1.96 × SEM × the square root of 2) was used to determine the magnitude of change that would exceed the threshold of measurement error at a 95% confidence level. For nonparametric data, Spearman rho was calculated to determine the strength of the monotonic relationship between the two instruments. For parametric data, Pearson´s r was used to determine the linear correlation between the two instruments. Interclass Correlation Coefficient (ICC) was calculated to determine the variability of measurements between the two instruments. Measurement error was estimated using the standard error of measurement (SEM) and limits of agreements (LoA) (Bland-Altman plot) was calculated to determine the agreement of measurements between the two instruments. An alpha level of 0.05 was defined for the statistical significance of all the tests.

Results

Twenty-nine participants (n = 12 female) were recruited. Between participants, age ranged from 18–51 years (mean 29, SD 9.4 years), height ranged from 1.6 –1.91 m (mean 1.75, SD 0.09 m), weight ranged from 51 –91.4 kg (mean 73, SD 10.5 kg) and BMI ranged from 19.9 – 28.4 BMI (mean 23.7, SD 2.2 BMI).

Prior to the definitive analyses, the data were checked for statistical outliers and scatter plots were derived to illustrate the closeness between the methods. Scatterplots of the HR measured by the GF during cycling at both 50W and 100W demonstrated several outliers where some GF measurements did not reflect HR accurately (Figure 2).

Mean values for both devices ± SD, mean differences ± SD, correlations, MDC_{95}, SEM and 95% LoA for all activities measured in bpm are summarized in Table 1. HR from the GF and PL correlated very highly during rest ($r = 0.997$ $p < .0001$), treadmill running at 8.7 km/h ($r = 0.906$ $p < .0001$) and rapid arm movements following 1 min. rest (Rho = 0.928 $p < .0001$). High correlations between GF and PL measures were found during cycling at 150 W (Rho = 0.889 $p < .0001$), treadmill running at 12.1 km/h ($r = 0.845$ $p < .0001$) and rapid arm movement following exercise ($r > 0.745$ $p < .001$). Low correlations between GF and PL measures were found during cycling at 100 W (Rho = 0.462 $p = .012$) and treadmill walking at 4.8 km/h (Rho >0.481 $p = .008$) and negligible correlations were found between GF and PL during cycling at 50 W (Rho = 0.269 $p = .158$) (Table 1). (Mukaka, 2012) Bland–Altman analysis revealed HR variability during most activities where some GF measurements did not reflect HR accurately (Figure 3). Bland-Altman plots also indicated that the GF generally underestimated all outcome measures compared to the PL, a tendency that was most evident during cycling (Figure 3).

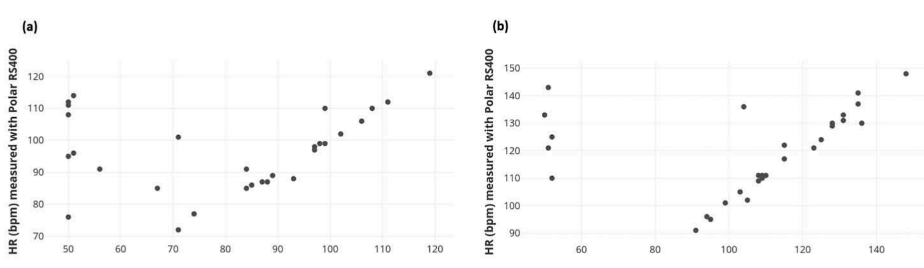

Figure 2. HR measured with the GF plotted against the HR measured by the PL during (a) cycling at 50W and (b) cycling at 100W.

Table 1. Mean, correlation, standard error of agreement between devices, MDC_{95} and Limit of agreement (Bland-Altman outcomes) for HR beats per minute (bpm) from the last measurement in each intensity level.

Intensity	Rest	Cycling 50 W	Cycling 100 W	Cycling 150 W	Treadmill 4.8 km/h	Treadmill 8.7 km/h	Treadmill 12.1 km/h	Rapid arm movement Following 1 min rest	Rapid arm movement Following 1 min exercise
Garmin Forerunner 235 (Mean ± SD)	59.5 ± 10.8	80.6 ± 22	104.9 ± 28.6	134.9 ± 30.9	109.8 ± 18.6	142.1 ± 17.9	165.8 ± 16.4	118.1 ± 18.1	128 ± 15
Polar RS400 (Mean ± SD)	60.5 ± 10.8	96,7 ± 12.5	119.8 ± 15.4	143.4 ± 18.9	101.1 ± 17.4	142.2 ± 19.0	168.3 ± 17.3	117.9 ± 18.4	145 ± 20.4
Mean Diff ± SD	1 ± 2.3	16.1 ± 22.9	14.9 ± 29.1	8.5 ± 22.6	8.7 ± 18.4	0.1 ± 8.1	2.5 ± 9.4	0.2 ± 7	17 ± 13.6
ICC (95% CI)	0.997 (0.968–0.994)	0.18 (−0.193–0.508)	0.195 (0.029–0.414)	0.609 (0.318–0.795)	0.480 (0.144–0.717)	0.905 (0.807–0.954)	0.844 (0.694–0.923)	0.928 (0.852–0.965)	0.71 (0.471–0.853)
r/Rho* (95% CI)	0.997 (0.944–0.996)	0.269* (−0.108 −0.578)	0.462* (−0.115–0.763)	0.889* (0. 776 −0.946)	0.481 (0.139 −0.72)	0.906 (0.808–0.955)	0.845 (0.694–0.925)	0.928* (0.851–0.965)	0.745 (0.521–0.873)
MDC95	0.3	57.5	72.4	44.4	34.0	8.8	14.5	5.2	20.3
SEM	0.1	20.7	26.1	16.0	13.2	3.2	5.2	1.9	7.3
Upper LOA	5.5	61.0	71.9	63.3	41.6	18.0	26.6	13.0	14.9
Lower LOA	−3.5	−28.8	−42.1	−37.1	−25.0	−22.4	−25.2	−14.4	−38.5

ICC = Intraclass Correlation Coefficient, CI = confidence interval, SD = standard deviation, Correlations are Pearson's correlation coefficient (r) except where indicated by * where they are Spearman rank correlation coefficients (Rho) due to non-normally distributed data.

Discussion

The present study investigated the concurrent validity of the wearable activity tracker Garmin Forerunner 235 during rest, cycling, treadmill walking and running, and rapid arm movements. The GF yielded HR values that closely correlated with the criterion measure PL, suggesting that GF-provided measures are valid for tracking HR changes during rest, treadmill running at 8.7 km/h and 12.1 km/h, cycling at 150W and rapid arm movements. This is consistent with previous studies assessing other PPG wearable devices (Parak & Korhonen, 2014; Stahl et al., 2016; Wallen et al., 2016). However, due to limited correlation, the results suggest that compared with the PL the GF did not provide valid measures for tracking changes in HR during treadmill walking at 4.8 km/h and cycling at 50W and 100W. This is in contrast to other studies which found a high correlation between PPG HR monitors and criterion measures during cycling and treadmill walking at similar speeds and intensities (Shcherbina et al., 2017; Stahl et al., 2016).

Present result indicates, that the GF generally underestimated HR compared with the criterion measure PL (Figure 3). This is supported by previous studies indicating that wearable PPG devices generally underestimate the true HR (Gillinov et al., 2017; Wallen et al., 2016). Further examination of the data revealed extreme variability around the average HR values measured by the GF especially during cycling at 50W and 100W (Figure 2). Upon closer inspection, several outliers were found to consistently measure HR values at exactly 50–51 bpm exclusively during cycling at 50W and 100W.

Several potential sources of error in PPG HR measures have been previously suggested, such as motion artefacts, misalignment between the skin and the optical sensor, variations in skin colour and poor tissue perfusion (Alzahrani et al., 2015). However, these factors could not account for the variability observed in the present study as the systematic outliers were found exclusively during ergometer cycling. Based on our interpretation of the data, this variability can be explained only by an unidentified systematic measurement error.

Previous studies have found that the accuracy of PPG measures increases with intensity levels (Shcherbina et al., 2017; Stahl et al., 2016) and it has been hypothesised that perfusion increases with increased intensity, which could decrease PPG measurement error (Stahl et al., 2016). The results of the present study show that the accuracy of the GF rose with intensity during cycling but not during treadmill running above 8.7 km/h (Table 1). Hence, this theory is partly supported. Increased accuracy during treadmill running has been found at speeds up to 9.6 Km/h (Stahl et al., 2016), but the present study is the first to examine HR at specific running speeds above 9.6 Km/h. The results of the present study show that the accuracy of the GF during treadmill running decreased slightly at 12.1 km/h compared with 8.7 km/h (Table 1). This reduction in PPG accuracy at higher running speed may be caused by increased measurement error due to movement artefacts following increases in body movements due to increased exercise intensity. (Spierer et al., 2015) Based on the present result, we suggest, that the PPG accuracy would further diminish at running speeds exceeding 12.1 km/h.

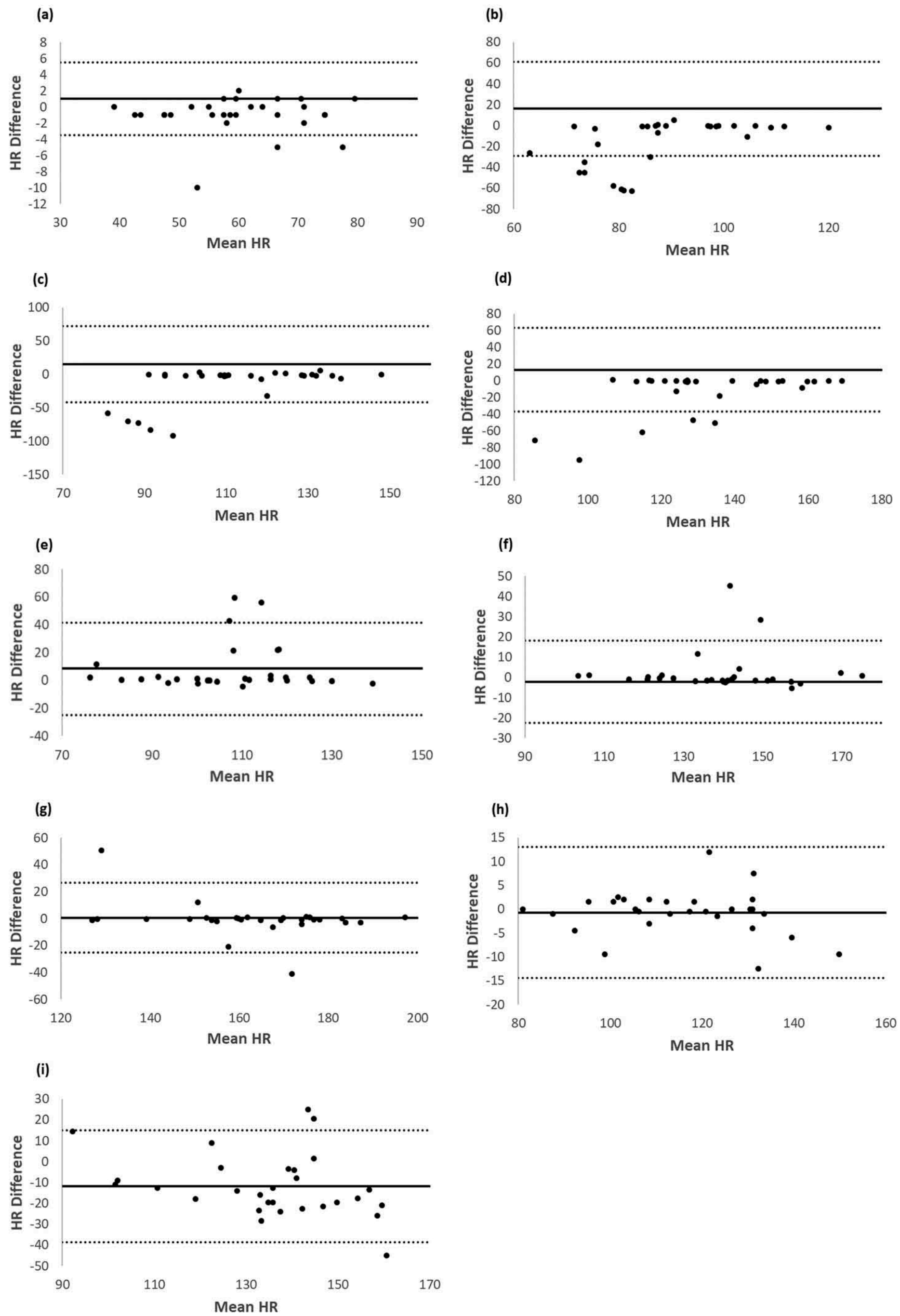

Figure 3. Bland–Altman plots and 95% limits of agreement with HR measured by the PL and GF during all activities and intensities. (a) Rest. (b) Cycling at 50W. (c) Cycling at 100W. (d) Cycling at 150W. (e) Walking at 4.1 km/h. (f) Running at 8.7 km/h. (g) Running at 12.1 km/h. (h) Rapid arm movement (following rest). (i) Rapid arm movement (following activity). Please take note of the different scaled y and x axes.

It has been proposed that changes in wrist position may render data collection less accurate due to changes in the placement of the PPG sensors (Spierer et al., 2015), and reduced accuracy of PPG HR measures has been linked to movement artefacts caused by arm movement (Gillinov et al., 2017). In the present study, we examined the measurement accuracy during high amplitude movements in multiple planes of motion using the rapid arm movement protocol as described by Spierer et al. (2015) (Spierer et al., 2015). The findings from the rapid arm protocol in the present study clearly showed a reduced accuracy of PPG HR measures during rapid arm movement (Figure 3). The results of the present study thus support the findings of a recent study, indicating that the accuracy of PPG HR measures is activity-dependent, with PPG technology performing best during activities with decreased arm movement (Gillinov et al., 2017).

In the present study, we observed that the GF registered changes in HR values with a considerable latency compared with the PL. This phenomenon, although most evident during rapid arm movements (Figure 3), was also observed during other physical activities (e.g., cycling, walking and running), and during both increases and decreases in HR. Movement artefacts could to some extent be accountable for this latency, at least during rapid arm movements and treadmill running at high speed. However, this does not explain the observations during cycling, where no arm movement occurred. To our knowledge, the issue of latency concerning PPG HR monitors has not previously been scientifically documented. Based on the present results further investigation concerning the magnitude and severity of PPG latency is warranted.

Considering the potential issue of latency, the application of wearable devices with PPG technology may yet be limited to monitoring HR during physical activities where HR reaches steady state or when data regarding ultimate performance is not required. This is supported by previous findings (Parak & Korhonen, 2014).

The results of the present study add to the existing literature on HR monitoring as this is the first study to undertake validation of a wearable PPG device at specific running speeds exceeding 9.6 Km/h. It is also the first study to observe considerable latency in wearable PPG measures, which may ultimately affect its application.

Despite the possible limitations, wearable devices with PPG HR sensors, however, do have the potential to overcome limitations of the traditional chest strap monitors (Stahl et al., 2016)and advance the performance of continuous HR measures at extended periods of time, as the reduced accuracy of PPG technologies is partially compensated by better usability and comfort (Parak & Korhonen, 2014).

Limitations

The study was conducted using a controlled protocol and transfer of results to free-living conditions should, therefore, be made with caution. The sample population included only healthy, younger individuals (18–51 years) with a BMI ranging from 19.9 – 28.4. Hence, generalisations cannot be made for youth and/or older adult age groups or for individuals of other body sizes or patient populations.

In accordance with current findings, previous studies have also documented the presence of outliers when measuring HR using PPG technology (Shcherbina et al., 2017; Weippert et al., 2010). The data collection procedures used in the present study were in accordance with manufacturer recommendations hence it is plausible that others may encounter similar findings using these procedures. Due to this, the observed outliers were included in the statistical analysis.

It is known that the occurrence of ectopic heart beats may affect HR measures, yet the presence of ectopic beats was not accounted for in the present study, hence this should be considered a limitation.

Conclusions

The present study showed that the Garmin Forerunner 235 yielded valid measures of HR during rest, treadmill running, rapid arm movement and cycling at 150W, although results varied across exercise conditions and intensities and measurement latency may potentially affect its application. Additional research is needed to assess wearable PPG HR variability at low intensity activities.

Acknowledgments

The authors wish to thank Maiken Morrel Lerche, BPt. for her contribution to the study.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the University College og Northern Denmark under Grant number [40334580].

ORCID

Morten Pallisgaard Støve http://orcid.org/0000-0002-6010-4376
Birgit Tine Larsen http://orcid.org/0000-0002-7829-6038

References

Alzahrani, A., Hu, S., Azorin-Peris, V., Barrett, L., Esliger, D., Hayes, M., … Kuoch, S. (2015). A multi-channel opto-electronic sensor to accurately monitor heart rate against motion artefact during exercise. *Sensors*, *15* (10), 25681–25702.

de Rezende Barbosa, M. P. D. C., Silva, N. T. D., de Azevedo, F. M., Pastre, C. M., & Vanderlei, L. C. M. (2016). Comparison of background: SpringGreen polar ® RS800G3 ™ heart rate monitor with background: SpringGreen 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*(2), 112–117.

Engström, E., Ottosson, E., Wohlfart, B., Grundström, N., & Wisén, A. (2012). Comparison of heart rate measured by Polar RS400 and ECG, validity and repeatability. *Advances in Physiotherapy*, *14*(3), 115–122.

Fallow, A. B., Tarumi, T., & Tanaka, H. (2013). Influence of skin type and wavelength on light wave reflectance. *Journal of Clinical Monitoring and Computing*, *27*, 313–317.

Forsberg, P., Analyseinstitut, I., Idræts-Forbund, D., & Forbund, D. A. (2012). Motionsløbere i Danmark : Portræt af danske motionsløbere. Idrættens Analyseinstitut. Retrieved from http://www.idan.dk/vidensbank/downloads/motionsloeb-i-danmark-portraet-af-de-danske-motionsloebere/b6326ae1-9d2c-4744-b5bf-9ff800a868fd

Gillinov, S., Etiwy, M., Wang, R., Blackburn, G., Phelan, D., Gillinov, A. M., … Desai, M. Y. (2017). Variable accuracy of wearable heart rate monitors during aerobic exercise. *Medicine and Science in Sports and Exercise*, *49* (8), 1697–1703.

Lee, J. M., Kim, Y., & Welk, G. J. (2014). Validity of consumer-based physical activity monitors. *Medicine and Science in Sports and Exercise*, *46*(9), 1840–1848.

Malisoux, L., Nielsen, R. O., Urhausen, A., & Theisen, D. (2015). A step towards understanding the mechanisms of running-related injuries. *Journal of Science and Medicine in Sport*, *18*(5), 523–528.

Mukaka, M. M. (2012). Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal*, *24* (3), 69–71.

Nunan, D., Gay, D., Jakovljevic, D. G., Hodges, L. D., Sandercock, G. R. H., & Brodie, D. A. (2009). Validity and reliability of short-term heart-rate variability from the Polar S810. *Medicine and Science in Sports and Exercise*, *41*(1), 243–250.

Parak, J., & Korhonen, I. (2014). Evaluation of wearable consumer heart rate monitors based on photopletysmography. *Conference Proceedings : … Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE Engineering in Medicine and Biology Society. Chicago: Annual Conference, 2014 (August 2014), 3670–3673. doi:10.1109/EMBC.2014.6944419

Shcherbina, A., Mattsson, C., Waggott, D., Salisbury, H., Christle, J., Hastie, T., … Ashley, E. (2017). Accuracy in Wrist-Worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort. *Journal of Personalized Medicine*, *7*(2), 3.

Spierer, D. K., Rossen, Z., Litman, L. L., & Fuji, K. (2015). Validation of photoplethysmography as a method to detect heart rate during rest and exercise. *Journal of Medical Engineering & Technology*, *39*, 264–271.

Stahl, S. E., An, H.-S., Dinkel, D. M., Noble, J. M., & Lee, J.-M. (2016). How accurate are the wrist-based heart rate monitors during walking and running activities? Are they accurate enough? *BMJ Open Sport & Exercise Medicine*, *2*(1), e000106.

Vasconcellos, F. V. A., Seabra, A., Cunha, F. A., Montenegro, R. A., Bouskela, E., & Farinatti, P. (2015). Heart rate variability assessment with fingertip photoplethysmography and poler RS800cx as compared with electrocardiography in obese adolescents. *Blood Pressure Monitoring*, *20*, 351–360.

Wallen, M. P., Gomersall, S. R., Keating, S. E., Wisløff, U., & Coombes, J. S. (2016). Accuracy of heart rate watches: Implications for weight management. *PLoS ONE*, *11*(5), 1–9.

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*, *109*(4), 779–786.