Comparison of Polar M600 Optical Heart Rate and ECG Heart Rate during Exercise

JOHN F. HORTON[1], PRO STERGIOU[1], TAK S. FUNG[2], and LARRY KATZ[3]

[1]Canadian Sport Institute Calgary, Calgary, Alberta, CANADA; [2]Research Computing Services, Information Technologies, University of Calgary, Calgary, Alberta, CANADA; and [3]Faculty of Kinesiology, University of Calgary, Calgary, Alberta, CANADA

ABSTRACT

HORTON, J. F., P. STERGIOU, T. S. FUNG, and L. KATZ. Comparison of Polar M600 Optical Heart Rate and ECG Heart Rate during Exercise. *Med. Sci. Sports Exerc.*, Vol. 49, No. 12, pp. 2600–2607, 2017. **Purpose**: The purpose of this study was to evaluate the accuracy of the Polar M600 optical heart rate (OHR) sensor compared with ECG heart rate (HR) measurement during various physical activities. **Methods**: Thirty-six subjects participated in a continuous 76-min testing session, which included rest, cycling warm-up, cycling intervals, circuit weight training, treadmill intervals, and recovery. HR was measured using a three-lead ECG configuration and a Polar M600 Sport Watch on the left wrist. Statistical analyses included OHR percent accuracy, mean difference, mean absolute error, Bland–Altman plots, and a repeated-measures generalized estimating equation design. OHR percent accuracy was calculated as the percentage of occurrences where OHR measurement was within and including ±5 bpm from the ECG HR value. **Results**: Of the four exercise phases performed, the highest OHR percent accuracy was found during cycle intervals (91.8%), and the lowest OHR percent accuracy occurred during circuit weight training (34.5%). OHR percent accuracy improved steadily within exercise transitions during cycle intervals to a maximum of 98.5% and during treadmill intervals to a maximum of 89.0%. Lags in HR calculated by the Polar M600 OHR sensor existed in comparison to ECG HR, when exercise intensity changed until steady state occurred. There was a tendency for OHR underestimation during intensity increases and overestimation during intensity decreases. No statistically significant interaction effect with device was found in this sample on the basis of sex, body mass index, $\dot{V}O_{2max}$, skin type, or wrist size. **Conclusions**: The Polar M600 was accurate during periods of steady-state cycling, walking, jogging, and running, but less accurate during some exercise intensity changes, which may be attributed to factors related to total peripheral resistance changes and pulse pressure. **Key Words:** PHOTOPLETHYSMOGRAPHY, ACCURACY, PHYSICAL ACTIVITY, STEADY STATE, BLAND–ALTMAN, ACTIVITY MONITOR

Advances in technology in recent years have allowed many companies developing wearable devices to offer activity monitors with sensors that measure heart rate (HR) at the wrist as an alternative to wearing an HR chest strap. HR measurement at the wrist is based on photoplethysmography (PPG) and is known as optical HR (OHR) measurement. PPG was introduced by Hertzman (8) in 1938 as a method for determining relative blood volume changes in the microvascular bed of peripheral tissues and as a methodology for evaluating peripheral circulation (6,13). In a review of PPG and its application in clinical physiological measurements, Allen (1) noted that PPG technology has been widely used in commercially available medical devices to measure blood pressure, cardiac output, blood oxygen saturation, and respiration and to assess peripheral vascular disease. The use of PPG is commonly accepted and can offer important diagnostic information on the cardiovascular system (1,6).

Although chest strap–based HR is determined by the R-R intervals in an ECG signal, PPG detects the frequency of blood pulses to measure HR at the wrist veins (1). OHR sensors placed against the skin typically use two optoelectronic components: a light-emitting diode (LED) and a photodiode (1). The LED illuminates the skin and the photodiode detects the intensity of the light reflecting and scattering back from the skin (13,21). The measured light intensity varies synchronously with the blood pulse. Common LED colors used in PPG sensors include green, red, and infrared; however, green LEDs are used in OHR sensors because they are known to produce the strongest plethysmographic signal for light reflectance measurement and detect pulse rate with a higher degree of precision (4,11,13,18). Currently, a two-LED solution is typically used by OHR sensor manufacturers, whereas some devices use three or more LEDs.

Wearable device reviewers and authors have complimented products with OHR sensors for their ease of use and comfort compared with traditional HR chest straps, but have acknowledged challenges with accuracy (4,7,9,10,12,14–16,18–20). Reasons for inaccurate OHR measurement include the

---

Address for correspondence: John F. Horton, M.Sc., Canadian Sport Institute Calgary, Room 125 Olympic Oval, 2500 University Drive NW, Calgary, Alberta, Canada T2N 1N4; E-mail: jhorton@csicalgary.ca.
Submitted for publication April 2017.
Accepted for publication July 2017.

0195-9131/17/4912-2600/0
MEDICINE & SCIENCE IN SPORTS & EXERCISE®

DOI: 10.1249/MSS.0000000000001388

following: weak peripheral blood circulation due to cold skin (4,21); distorted optical pulse signal resulting from motion artifact (4,10,12,21); improper device attachment where the wristband is too tight, too loose, or poorly positioned (14,19–21); skin type (4,10,18); sensor location on the body (10); and wrist position during activity (18). Reports by sport product reviewers and scientists have noted inaccurate HR measurement at the beginning of exercise, at high intensities, during interval training, and while speed changes during locomotion (9,10,16,18,20). There have also been reports of different degrees of accuracy depending on the activity such as cycling and weight lifting, which involve gripping with fingers and/or flexion/extension of the wrist (18). Although consumer reviews of OHR products offer valuable, real-world assessments, their conclusions are based on personal experiences and testing sessions that lack proper scientific design, which can lead to misinterpretation. Several studies have been conducted to investigate OHR accuracy during various activities (4,7,9,14,15,17–19,23); however, fewer have compared with ECG HR as a gold standard reference (7,9,14,15,17,23).

Polar Electro recently introduced the M600 Sport Watch to the market with a six-LED OHR sensor, which provides more optical measurement channels compared with other LED configurations. The six-LED OHR sensor offers more options to select and use the optimal OHR signal and is believed to improve overall accuracy by enhancing motion artifact detection and rejection and to improve OHR quality when the device is worn too loose or too tight. The purpose of this study was to evaluate the accuracy of the Polar M600 OHR measurement compared with ECG HR measurement at rest, during various physical activities, and during recovery.

METHODS

Subjects

Individuals between 18 and 55 yr and between a body mass index (BMI) of 20–27 kg·m^{-2} were included in this study. Thirty-six men and women (age, 40.5 ± 9.6 yr; height, 172.6 ± 9.2 cm; weight, 69.8 ± 11.0 kg; mean ± SD) volunteered to participate after completing a Physical Activity Readiness Questionnaire (PAR-Q+) and giving written informed consent. This study was approved by the Conjoint Health Research Ethics Board of the University of Calgary. Detailed demographic information is summarized in Table 1.

Skin Type Assessment

Skin type was assessed using the Fitzpatrick Skin Type Scale (5). Subjects were provided a document outlining the 6-point scale with example photographs and characteristics and asked to self-assess their skin type. Type 1 is considered ivory, whereas type 6 is considered very dark brown.

TABLE 1. Subject demographics and protocol information.

	Male (*n* = 18)	Female (*n* = 18)
Demographic information		
Age, yr	39.3 ± 9.6	41.8 ± 9.7
Height, cm	179.3 ± 6.0	166.0 ± 6.6
Weight, kg	78.1 ± 7.9	61.6 ± 6.8
BMI, kg·m^{-2}	24.3 ± 2.3	22.3 ± 2.0
Skin type, 1–6	2.7 ± 0.9	2.3 ± 0.6
Estimated $\dot{V}O_{2max}$, mL·kg^{-1}·min^{-1}	43.9 ± 6.4	44.3 ± 4.4
Right wrist circumference, cm	17.0 ± 1.0	15.1 ± 0.8
Left wrist circumference, cm	16.9 ± 1.0	15.0 ± 0.8
Protocol details		
Dumbbell weight, kg	8.1 ± 1.3	5.8 ± 1.0
Cycle warm-up workload, W	109.7 ± 24.5	81.9 ± 14.4
Cycle intervals workload		
Workload 1, W	119.4 ± 25.1	80.6 ± 20.2
Workload 2, W	144.4 ± 25.1	105.6 ± 20.2
Workload 3, W	169.4 ± 25.1	130.6 ± 20.2
Workload 4, W	194.4 ± 25.1	155.6 ± 20.2
Treadmill intervals speed		
Walk, km·h^{-1}	4.0 ± 0.0	4.0 ± 0.0
Jog, km·h^{-1}	8.0 ± 0.0	8.0 ± 0.0
Run, km·h^{-1}	12.2 ± 1.1	11.6 ± 1.1

All values are mean ± SD.

Estimated $\dot{V}O_{2max}$

Estimated $\dot{V}O_{2max}$ was calculated by using HR and workload data from the first four stages of a seven-stage pyramid-style cycle interval protocol on a Velotron cycle ergometer (RacerMate Inc., Seattle, WA). Stages were 3 min long and increased in intervals of 25 W. Mean ECG HR from the last 20 s of each stage was recorded. HR and corresponding workload for each stage were used in an equation to calculate the slope of the line of best fit between HR and workload, and then used to extrapolate to estimated maximal HR (220 − age) and interpolate maximal power output. $\dot{V}O_{2max}$ was calculated as a measure of cardiovascular fitness using the American College of Sport Medicine $\dot{V}O_{2max}$ metabolic estimation equation (2) for leg cycling ($\dot{V}O_{2max}$= max W·kg^{-1} × 10.8 + 7).

Experimental Protocol

Each subject performed designated activities in the same order during a continuous 76-min session in a controlled laboratory environment (mean temperature, 20°C; mean humidity, 25%). The entire testing session is herein referred to as all activities and included rest, exercise recovery, cycle warm-up, cycle intervals, circuit weight training, treadmill intervals, and activity change periods. Graphical representation of the testing session is found in Figure 1, and protocol details are presented in Table 1. *Rest*: Subjects sat calmly in a chair for 5 min with arms on arm rest. No talking was permitted. *Cycle warm-up*: Subjects performed a 5-min cycle warm-up on a Velotron cycle ergometer. A light workload (e.g., 50–100 W) was chosen for each subject and HR was monitored throughout. The cycle warm-up was used to raise body temperature before more intense exercise and to assess HR response to the selected workload, which then assisted in selecting the initial workload during cycle intervals. *Cycle intervals*: Subjects performed 21 min of intervals in a pyramid fashion on a Velotron cycle ergometer.

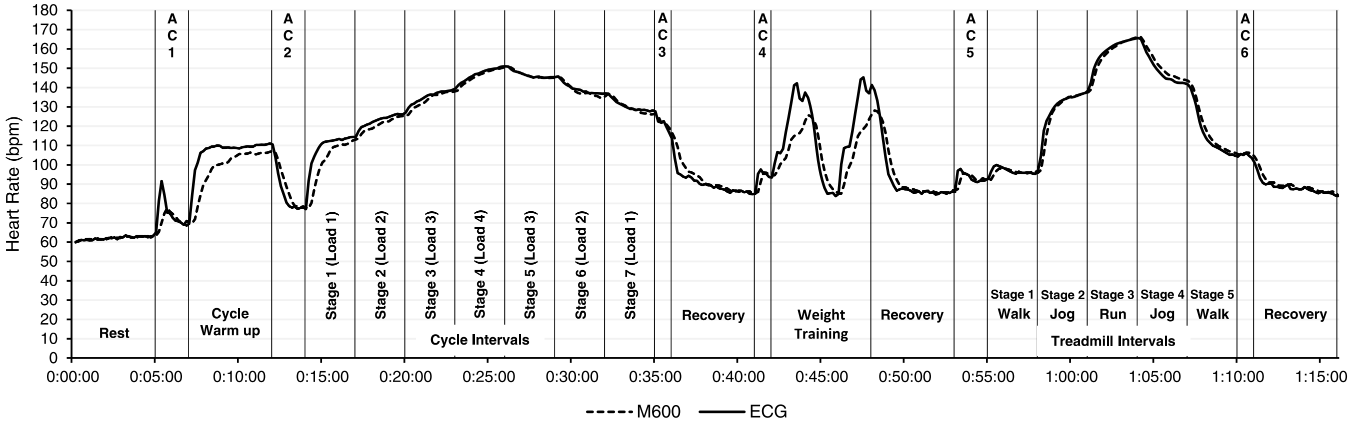

FIGURE 1—A 76-min testing session graph showing exercise phases, time periods for activities, and comparison of Polar M600 HR and ECG HR. All 10-s mean HR samples for all subjects. AC, activity change periods.

Stages were 3 min in duration and changed by 25 W (e.g., workload 1, 100 W; workload 2, 125 W; workload 3, 150 W; workload 4, 175 W; workload 3, 150 W; workload 2, 125 W; workload 1, 100 W). Workload 1 was chosen for each subject on the basis of HR response throughout the 5-min cycle warm-up and was typically between 105 and 115 bpm. *Circuit weight training*: Subjects performed a circuit weight training session with a dumbbell in each hand. Mean weight used during the circuit weight training protocol by male and female subjects was 8.1 ± 1.3 and 5.8 ± 1.0 kg, respectively. Exercises included shoulder shrugs, squats, bicep curls, and lunges. Each exercise was performed for 30 s with no rest between exercises. After the four-exercise circuit, subjects placed the dumbbells on the floor and sat in a chair for a 2-min rest period and then repeated the circuit. *Treadmill intervals*: Subjects performed a 15-min pyramid format interval session with five, 3-min stages on a treadmill (Woodway 4Front View Treadmill, Waukesha, WI) at 0° incline in the following manner: walk, jog, run, jog, and walk. Walking speed was 4.0 km·h^{-1} and jogging speed was 8.0 km·h^{-1}. The running speed was selected by each subject on the basis of recent 5-km race pace (male, 12.2 ± 1.1 km·h^{-1}; female, 11.6 ± 1.1 km·h^{-1}; mean ± SD). *Recovery*: Subjects sat in a chair for 5 min after cycle intervals, circuit weight training, and treadmill intervals. Talking was permitted. *Activity change*: Throughout the 76-min session, six activity change periods (1 or 2 min) were included to allow subjects time to move from one location to another in the laboratory testing area or to prepare for exercise phases.

Data Collection

HR data were collected with a three-lead ECG configuration using Power Lab 16/30 with Bio Amp model ML132 and Lab Chart Pro 7.1 Software (AD Instruments, Castle Hill, Australia), and a Polar M600 Sport Watch (Polar Electro Oy, Kempele, Finland). Skin was cleaned with alcohol and AgAgCl surface electrodes with a 19 × 16-mm active area (3M Red Dot Monitoring Electrode 2560; 3M Health Care, St. Paul, MN) were placed on the skin at V2, V6, and clavicle. As recommended by the manufacturer in the instruction manual, the Polar M600 was secured to the left lower arm just above the ulnar styloid process and tight enough to prevent device movement but not uncomfortable for the subject. ECG HR data collection was initiated using Lab Chart Pro 7.1 and started first with a 30-s lead-in before starting the Polar M600. At the 25-s mark, a researcher began a 5-s countdown to start the Polar M600. Several minutes before the beginning of data collection, the Polar M600 was switched into Other Indoor training mode where HR values were displayed on the watch face once acquired by the device. This process was done for all subjects to ensure that HR was being measured before starting the testing session. Other Indoor training mode was used for all subjects and for all activities. A "start" marker was inserted in Lab Chart Pro 7.1 to be used later for synchronization of ECG and Polar M600 HR data. Three different Polar M600 devices were equally used throughout the study.

Data Analysis

Data processing. ECG data were sampled at 1000 Hz to display the PQRST waveform in Lab Chart Pro 7.1. Using an algorithm in Lab Chart Pro 7.1, HR was calculated from the time between the R-R intervals. ECG HR data were then down-sampled from 1000 Hz and exported as a text file at 1-s intervals. The training session data file from the Polar M600 was synched to a mobile device with the Polar Flow App and then HR data were downloaded at 1-s intervals using the Polar Flow Web service. The two HR data files for each subject were synchronized by using the start marker in the ECG data file and the first Polar M600 HR sample. Every 10 s throughout the data files, mean HR was calculated for both measurement devices. Markers inserted within the ECG HR data in Lab Chart Pro 7.1 to identify event start times were used to separate the HR data for various activities throughout the 76-min testing session.

Statistical analyses. OHR percent accuracy was defined and calculated as the percentage of occurrences where the OHR measurement was within and including ±5 bpm from the ECG HR value, a range similarly used by others (4). Mean difference was calculated as the mean of the differences between OHR measurement and ECG HR measurement either in a positive or negative direction. A negative value represented underestimation by the Polar M600 compared with ECG measurement. Mean absolute error (MAE) was calculated as the mean of absolute differences between the OHR measurement and ECG HR measurement. Agreement between the Polar M600 HR and ECG HR was analyzed using Bland–Altman plots (3), where mean difference, SD, SD × 1.96, and upper and lower limits of agreement (LOA) were calculated.

Data were further analyzed using an unbalanced repeated-measures design with interval outcome variables using a generalized estimating equation (GEE; i.e., GEE under GENLIN procedure in SPSS v.22) to determine the effects of the Polar M600 HR and ECG HR on outcomes. The data from this study met the basic assumptions of GEE analyses: the distribution was normal and considered the correlated nature of observations within each participant and unequal number of observations among participants. Also, the model allowed for the addition of predictors including device, sex, BMI, $\dot{V}O_{2\max}$, skin type, and wrist circumference. An α level of 0.05 was chosen, and any computed P values less than 0.05 were considered statistically significant.

Sample size calculation. Sample size calculation was performed using a method for comparison of a mean to a known value (inference of a mean). MAE and SD from pilot data collected for exercise protocols of walking, jogging, and running were used as input criteria for the sample size calculation (difference between means, 4 bpm; SD of this measurement, 6 bpm). Using an α of 0.05 and a power of 0.80, the sample size was calculated to be 36 subjects.

RESULTS

OHR percent accuracy, mean difference, MAE, and upper and lower LOA values for the various activities performed throughout the testing session when comparing Polar M600 and ECG measured HR are summarized in Table 2. Results in this table are ranked from the highest to the lowest OHR percent accuracy. Results for each stage during cycle intervals and treadmill intervals are shown as subheadings and are arranged in a stage order. Of the four exercise protocols performed, the highest OHR percent accuracy was found during cycle intervals (91.8%), whereas the lowest OHR percent accuracy occurred during circuit weight training (34.5%). During cycle intervals, accuracy improved consistently with exercise transition until stage 5, where OHR percent accuracy reached 98.5% and remained greater than 95% for the remainder of this interval session. Accuracy also improved with exercise transition through the first four stages during treadmill intervals, where the highest OHR percent accuracy (89.0%) was found during the second jog segment (stage 4). In addition, OHR percent accuracy during rest was 93.5% with a mean difference of −0.1 bpm and an MAE of 1.9 bpm.

Lags in HR measured by the Polar M600 compared with the ECG-measured HR were observed during the 76-min testing session during different activities while HR was increasing or decreasing (Fig. 1). HR lag was defined as an OHR underestimation during exercise intensity increase or an OHR overestimation during exercise intensity decrease, or simply an OHR under/overestimation for a duration until HR stabilization occurred relative to ECG HR measurement. HR lag was observed during the first 2 min of cycle warm-up and during stage 1 of cycle intervals, circuit weight training, and speed changes throughout treadmill intervals (Fig. 1).

Bland–Altman plot analyses showed the distribution of error with respect to the mean difference between the Polar

TABLE 2. Summary of Polar M600 OHR accuracy.

	n	OHR % Accuracy	Mean Difference, bpm	SD, bpm	Mean Absolute Error, bpm	Upper LOA, bpm	Lower LOA, bpm
Rest	1080	93.5	−0.1	2.7	1.9	5.3	−5.4
Cycle intervals	4536	91.8	−1.9	7.0	2.8	11.8	−15.6
Stage 1	648	68.8	−7.3	12.6	7.9	17.3	−31.9
Stage 2	648	90.9	−2.4	6.8	2.9	11.0	−15.8
Stage 3	648	94.9	−1.6	5.5	2.1	9.1	−12.4
Stage 4	648	97.7	−1.0	1.9	1.3	2.6	−4.7
Stage 5	648	98.5	0.0	1.6	1.1	3.1	−3.1
Stage 6	648	95.1	−0.7	5.7	2.0	10.5	−11.9
Stage 7	648	96.6	−0.3	5.7	1.9	10.8	−11.5
Treadmill intervals	3240	81.0	0.3	5.6	3.2	11.3	−10.7
Stage 1	648	74.4	−0.5	5.1	3.6	9.5	−10.5
Stage 2	648	78.7	−1.9	6.0	3.7	9.9	−13.8
Stage 3	648	84.6	−1.3	4.5	2.7	7.4	−10.1
Stage 4	648	89.0	2.5	6.5	3.0	15.2	−10.1
Stage 5	648	78.5	2.7	3.6	3.2	9.8	−4.3
All activities	16,416	76.6	−1.4	9.5	4.8	17.2	−20.1
Recovery	3240	74.4	1.8	6.9	4.2	15.4	−11.8
Cycle warm-up	1080	70.5	−8.7	16.1	9.8	22.9	−40.4
Activity change	1944	61.3	0.0	8.6	5.8	16.8	−16.9
Weight training	1728	34.5	−6.2	17.0	12.7	27.2	−39.6

OHR percent accuracy is the percentage of occurrences where the OHR measured HR was within and including ±5 bpm from the ECG HR value. Results are ranked from the highest to the lowest OHR percent accuracy. Results for each stage during cycle intervals and treadmill intervals are shown as subheadings and are arranged in stage order.
n, number of samples on all subjects.

M600 and the ECG-measured HR (Figs. 2 and 3). Among the four exercise phases, the narrowest 95% LOA values existed in the treadmill interval data (upper LOA, 11.3 bpm; lower LOA, −10.7 bpm), and the widest 95% LOA values were found in the circuit weight training data (upper LOA, 27.2 bpm; lower LOA, −39.6 bpm). Narrow 95% LOA values also existed in the resting HR data (upper LOA, 5.3 bpm; lower LOA, −5.4 bpm).

Mean HR of all subjects measured by the Polar M600 and ECG for each activity and statistical significance from a GEE analysis are found in Table 3. A statistically significant difference was not found between the Polar M600 HR and the ECG HR for rest, treadmill intervals, and activity change ($P = 0.547$, $P = 0.496$, and $P = 0.920$, respectively). A statistically significant difference was found between the Polar M600 HR and the ECG HR for all other individual activities ($P < 0.001$).

GEE results showed that there was no statistically significant sex by device interaction effect ($\chi^2(1) = 0.083$, $P = 0.773$) and no statistically significant sex effect ($\chi^2(1) = 0.297$, $P = 0.586$). There was no statistically significant BMI by device interaction effect ($\chi^2(1) = 2.716$, $P = 0.099$) and no statistically significant BMI effect ($\chi^2(1) = 0.527$, $P = 0.468$). There was no statistically significant skin type by device interaction effect ($\chi^2(1) = 0.033$, $P = 0.857$) and no statistically significant skin type effect ($\chi^2(1) = 0.184$, $P = 0.668$). There was no statistically significant wrist circumference by device interaction effect ($\chi^2(1) = 1.173$, $P = 0.279$) and no statistically significant wrist circumference effect ($\chi^2(1) = 0.277$, $P = 0.598$). There was no statistically significant $\dot{V}O_{2max}$ by device interaction effect ($\chi^2(1) = 0.725$, $P = 0.394$), but there was a statistically significant $\dot{V}O_{2max}$ effect ($\chi^2(1) = 16.483$, $P < 0.001$) with estimated coefficient ($B = -1.271$, SE = 0.313), which means that $\dot{V}O_{2max}$ is negative and significantly affects the HR. In other words, HR is lower for subjects with a higher $\dot{V}O_{2max}$.

DISCUSSION

The present study compared the Polar M600 OHR measurement with a gold-standard ECG reference HR measurement during rest, cycle warm-up, cycle intervals, circuit weight training, treadmill intervals, and recovery. The highest OHR percent accuracy among the four exercise phases existed during cycle intervals. Furthermore, OHR percent accuracy improved steadily with exercise transitions during cycle intervals as HR increased to peak exercise values for this cycling activity. Accuracy also improved as HR increased throughout exercise transitions during treadmill intervals. These findings contradict conclusions from others (9) who found that OHR accuracy became worse during activities eliciting higher HR where the modes of exercise were treadmill jogging and running, and stair climbing.

Although cycle intervals resulted in a higher OHR percent accuracy (91.8%) than treadmill intervals (81.0%), a marginally worse mean difference was found (cycle intervals, −1.9 bpm; treadmill intervals, 0.3 bpm; Table 2). This finding may be explained by the initial OHR underestimation during stage 1 of cycle intervals, and both underestimation and overestimation

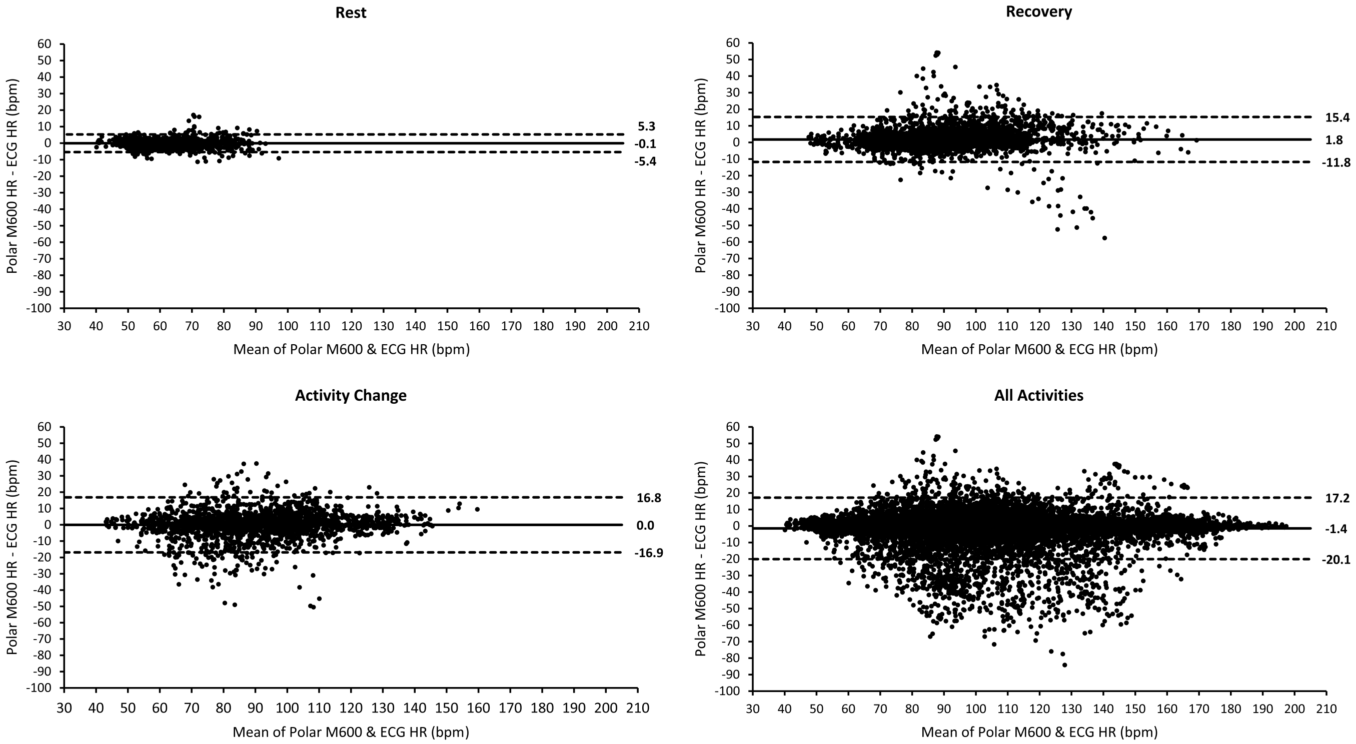

FIGURE 2—Bland–Altman plots showing the agreement of Polar M600 HR and ECG HR during rest, recovery, activity change, and all activities combined. All 10-s mean HR samples for all subjects.

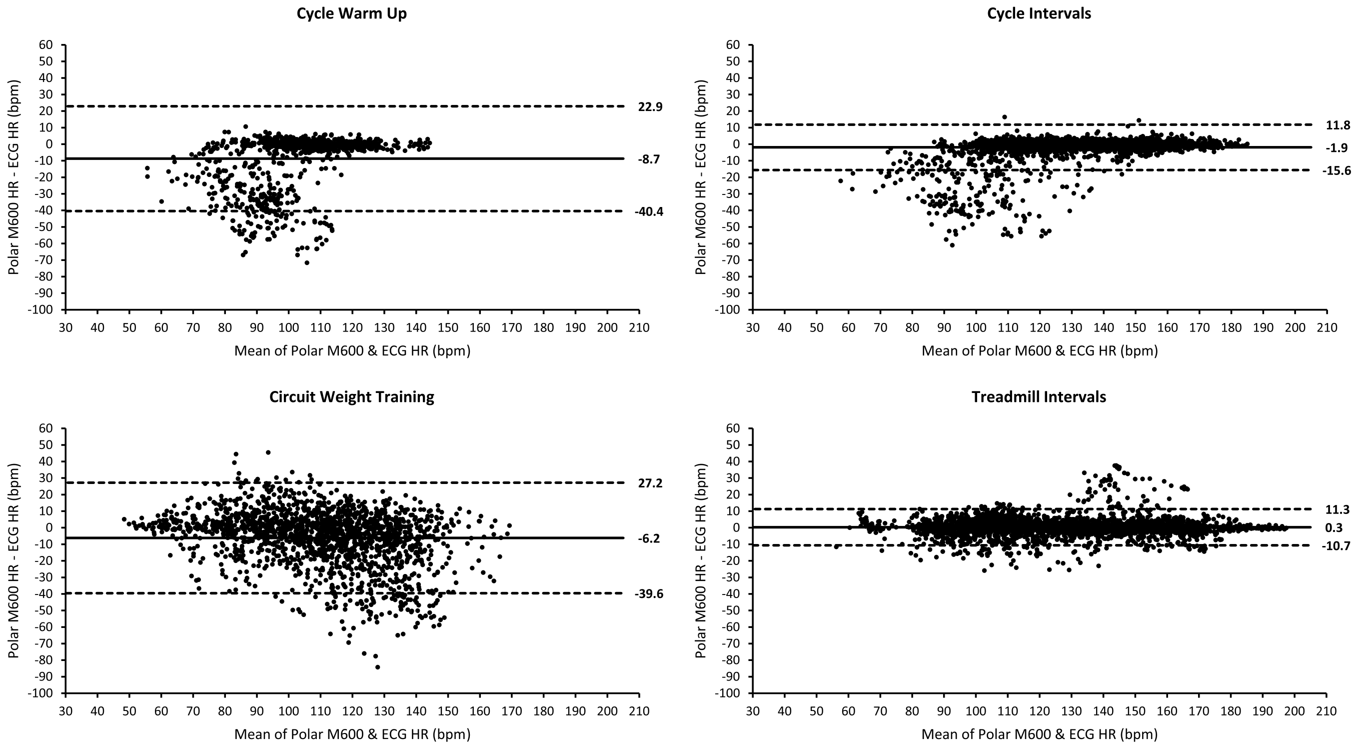

FIGURE 3—Bland–Altman plots showing the agreement of Polar M600 HR and ECG HR during cycle warm-up, cycle intervals, circuit weight training, and treadmill intervals. All 10-s mean HR samples for all subjects.

during different segments of treadmill intervals, which balanced the direction of error. This was also illustrated in Figure 3, where there was improved agreement and tighter grouping of data around the mean difference and within the LOA values for cycle intervals compared with treadmill intervals, except for the scattered data below the lower LOA, which mostly represents the OHR underestimation during cycle interval stage 1. It is also worth noting that most of the scattered data below the lower LOA (−15.6 bpm) in the cycle intervals using the Bland–Altman plot were a result of four subjects who exhibited periods of considerable OHR underestimation, and most of the scatter above the upper LOA (11.3 bpm) in the treadmill intervals using the Bland–Altman plot was a result of one subject's overestimated data (Fig. 3). Data from each of these outliers were examined thoroughly to determine if any anatomical or physiological variables, or specific observations during testing might account for the considerable OHR inaccuracies. The authors found no explanations for these OHR anomalies. It is possible that squeezing the handle bars on the cycle ergometer during cycle intervals and movement artifact during treadmill intervals contributed to the considerable OHR underestimation and overestimation among these subjects; however, this cannot be confirmed because grip pressure and arm movement were not measured in this study.

The most noticeable OHR inaccuracies were observed as lags in HR at the beginning of each exercise phase, during exercise transition, and throughout circuit weight training, which supports findings by others (16,18,20). A considerable HR lag occurred at the beginning of cycle warm-up. Similarly, lags in HR occurred during circuit weight training, which began 42 min into the testing session, and throughout treadmill intervals, which began 55 min into the testing session. Therefore, it cannot be concluded that the HR lag phenomenon only occurs at the beginning of an exercise session with an initial HR rise.

Other than stage 1 cycle interval, there were no considerable lags in HR during exercise transition throughout this cycle protocol (Fig. 1). In contrast, there were more HR lags during exercise transition throughout treadmill intervals. While increasing speed from stage 1 (walk) to stage 2 (jog), HR lagged for 40 s before OHR percent accuracy reached greater than 90% for the remainder of the first jog stage. A further 30-s HR lag was recorded between stages 2 (jog) and 3 (run) before OHR percent accuracy reached 90%. During speed decreases from stages 3 through 5 (run, jog, walk), there was an overestimation of the OHR measurement compared

TABLE 3. Overall HR summary of Polar M600 HR and ECG HR with statistical significance.

	HR, bpm		
	ECG	**Polar M600**	**Sig.**
Rest (*n* = 1080)	62.1 ± 1.8	62.1 ± 1.8	*P* = 0.547
Recovery (*n* = 3240)	90.1 ± 2.2	91.9 ± 2.1	*P* < 0.001*
Cycle warm-up (*n* = 1080)	107.8 ± 2.0	99.1 ± 2.3	*P* < 0.001*
Cycle intervals (*n* = 4536)	133.3 ± 2.0	131.3 ± 2.0	*P* < 0.001*
Circuit weight training (*n* = 1728)	113.2 ± 2.5	107.1 ± 2.2	*P* < 0.001*
Treadmill intervals (*n* = 3240)	130.0 ± 2.1	130.3 ± 2.1	*P* = 0.496
Activity change (*n* = 1944)	91.5 ± 1.9	91.4 ± 1.9	*P* = 0.920

Values are mean ± SEM of all 10-s mean HR samples for all subjects.
**P* < 0.05.

with ECG HR measurement for the entire duration by approximately 3 bpm (Fig. 1). Perhaps OHR percent accuracy was higher during cycle intervals than during treadmill intervals because the exercise transitions were subtler while performing cycling intervals. During the cycle intervals, HR increased by approximately 12 bpm with each exercise transition compared with 35 bpm during treadmill intervals. Similarly, the HR drop with exercise transition during cycle intervals was approximately 8 bpm compared with 30 bpm on the treadmill. Therefore, higher OHR percent accuracy exhibited during the cycle intervals may have occurred because the exercise intensity changes were not as great for this pyramid-style interval session.

The Polar M600 was least accurate during circuit weight training, with an OHR percent accuracy of 34.5%, a mean difference of −6.2 bpm, and an MAE of 12.7 bpm. There was very little agreement from a Bland–Altman plot analysis with wide LOA values (upper LOA, 27.2 bpm; lower LOA, −39.6 bpm; Fig. 3) and repeated lags in HR during the four-exercise circuit performed two times with rest periods (Fig. 1). In addition, HR measured by the Polar M600 did not reach peak HR values measured by ECG in either weight training circuit. Other researchers (18) have suggested that motion artifact from arm movement or wrist articulation may cause inaccuracies from OHR sensors. Limited arm movement and wrist articulation occurred during weight training in the present study where shoulder shrugs, squats, and lunges were all performed with arms by the side of the body. The only exercise in this study that involved considerable arm movement was bicep curls; therefore, it is speculated that OHR inaccuracies were a result of occlusion at the wrist (squeezing) while holding light dumbbells in the hands and not necessarily a result of motion artifact during the circuit weight training protocol.

GEE results showed that there was no statistically significant interaction effect with device in this sample on the basis of sex, BMI, $\dot{V}O_{2max}$, skin type, or wrist circumference. Few studies have tested for interaction effects of any of the above variables on OHR error (7,18,19). Findings by other researchers also found no significant effect on OHR error on the basis of sex (7,19), BMI (7), and wrist circumference (7); however, Spierer et al. (18) did find a significant effect on OHR error on the basis of skin type and concluded that less photosensitive skin types (i.e., darker skin) may result in greater OHR error. Participants in this study had a mean skin type of 2.5 ± 0.8 (range, 1–5), whereas the mean skin type of subjects in the study by Spierer et al. (18) ranged from 2 to 5 on the Fitzpatrick scale. Skin type is assumed to be similar between the two studies, but perhaps a larger study sample by Spierer et al. (18) compared with this study (50 vs 36), or the inclusion of more subjects with skin type 5, contributed to conflicting findings regarding skin type effect.

Although OHR sensor pressure against the skin was not measured in this study, device attachment was carefully monitored on all subjects to ensure proper placement and tightness. The Polar M600 fit larger wrist sizes better than smaller wrist sizes. This was noticed by the researchers and by several subjects with smaller wrists. Once the Polar M600 was tightened and secured with the strap clasps, there was a tendency for the chassis to roll slightly to the medial side (toward the fifth finger) of the wrist and not remain in the original centered position for subjects with smaller wrists. No statistically significant wrist circumference effect was found from GEE analysis; therefore, the authors speculate that OHR inaccuracies found in the present study were likely not a result of improper device attachment.

The Polar M600 exhibited greater accuracy as HR became higher during cardiovascular exercise and while at steady state. However, there were situations where the OHR sensor responded slowly in calculating HR response to intensity changes until system stabilization occurred. Physiological responses at the beginning of exercise, whether it is the first bout or subsequent bouts after recovery periods, are quickly detected by HR chest straps and ECG through electrical activity emitting from the heart. It seems that this OHR technology at the wrist is less sensitive to physiological responses to intensity changes during exercise. This may be explained by decreasing total peripheral resistance that obscures the potential change in pulse pressure (22), therefore affecting the detection of blood pulse and contributing to a lag time. The result is a delay in calculated HR response for periods of time until the body stabilizes to the new intensity. Scientists and consumers should consider that there is no issue with this OHR measurement in the periphery and that this sensor is measuring what it is designed to measure despite these HR lags. Although only the Polar M600 was tested during this study, the authors have conducted previous unpublished studies using similar testing protocols and sample sizes while using various brands of activity monitors with OHR sensors and found comparable results with respect to HR lag during intensity changes. It is hypothesized that wrist-based OHR sensors exhibit some degree of HR lag for reasons described previously. It is important to understand that this technology currently is not as accurate as ECG-based HR measurement and should be recognized with realistic acceptable error.

CONCLUSIONS

The Polar M600 was accurate during periods of steady-state cycling, walking, jogging, and running, but HR lags existed during intensity changes until system stabilization occurred. During weight training, reliance on OHR measurement should be used more so for monitoring HR recovery after lifting weights. Statistically significant differences were found between OHR and ECG HR measurements in this study. Although there is merit in determining statistically significant differences, one must also evaluate real-world applications and decide on an acceptable amount of error for their needs. For example, a statistically significant difference was found between OHR and ECG HR measurements for an activity in this study, but mean HR differed by only 2 bpm. It should be noted that this study sample was homogenous with respect to BMI; therefore, conclusions regarding the effects of BMI on OHR accuracy are limited to those individuals with a BMI between 20 and 27 kg·m^{-2}. Furthermore, the results of this study may not apply to older, less athletic individuals. Other OHR sensors may not have the same degree of accuracy as found with the Polar M600 in this study.

Future studies should investigate the effect of varying changes in exercise intensity, randomize the order of activities, include a heterogenous sample, and compare OHR measurement during lower-body and upper-body weight training exercises. Measuring movement of the lower arm while exercising would be valuable to assess the relationship of motion artifact and OHR accuracy, as well as investigating how proper device attachment (e.g., correct tightness and location on the wrist) affects OHR accuracy.

The authors thank Dr. David Smith for his exercise physiology expertise and editing, Sven Wörner for his dedicated assistance during data collection, Friedrich Czarnecki for participating in many pilot testing sessions, and the participants of this study.

Polar Electro Oy funded this research study and provided products for testing.

The Canadian Sport Institute Calgary conducted this study independently and unbiased for Polar Electro Oy. Results may benefit Polar Electro Oy.

The authors declare no conflicts of interest.

The results of this study do not constitute endorsement by the American College of Sports Medicine.

The results of this study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

REFERENCES

1. Allen J. Photoplethysmography and its application in clinical physiological measurement. *Physiol Meas*. 2007;28:R1–39.
2. American College of Sports Medicine. *ACSM's Guidelines for Exercise Testing and Prescription*. 7th ed. Philadelphia (PA): Lippincott Williams & Wilkins; 2006. p. 289.
3. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 372(8476):307–10.
4. Delgado-Gonzalo R, Parak J, Tarniceriu A, Renevey P, Bertschi M, Korhonen I. Evaluation of accuracy and reliability of PulseOn optical heart rate monitoring device. *Conf Proc IEEE Eng Med Biol Soc*. 2015:430–3.
5. Fitzpatrick TB. "Soleil et peau" [Sun and skin] [in French]. *J Méd Esth*. 1975(2):33–4.
6. Gil E, Orini M, Bailón R, Vergara JM, Mainardi L, Laguna P. Photoplethysmography pulse rate variability as a surrogate measurement of heart rate variability during non-stationary conditions. *Physiol Meas*. 2010;31(9):1271–90.
7. Gillinov S, Etiwy M, Wang R, et al. Variable accuracy of wearable heart rate monitors during aerobic exercise. *Med Sci Sports Exerc*. 2017;49(8):1697–703.
8. Hertzman AB. The blood supply of various skin areas as estimated by the photo-electric plethysmograph. *Am J Physiol*. 1938;124:328–40.
9. Jo E, Dolezal BA. *Validation of the Fitbit® Surge™ and Charge HR ™ Fitness Trackers*. [cited 1 Mar 2017]. Available from: http://www.lieffcabraser.com/pdf/Fitbit_Validation_Study.pdf.
10. LeBoeuf S. *Optical Heart-Rate Measurement's Top 5 Challenges*. [cited 1 Mar 2017]. Available from: http://www.edn.com/electronics-blogs/about-embedded/4440217/Optical-heart-rate-measurement-s-top-5-challenges.
11. Lee S, Shin H, Hahm C. Effective PPG sensor placement for reflected red and green light, and infrared wristband-type photoplethysmography. In: *18th International Conference on Advanced Communication Technology (ICACT)*. PyeongChang (South Korea): IEEE. 2016. pp. 556–58.
12. Lin Z, Zhang J, Chen Y, Zhang Q. Heart rate estimation using wrist-acquired photoplethysmography under different types of daily life motion artifact. In: *2015 IEEE International Conference on Communications (ICC)*. London (UK): IEEE; 2015. pp. 489–94.
13. Maeda Y, Sekine M, Tamura T. The advantages of wearable green reflected photoplethysmography. *J Med Syst*. 2010;35(5): 829–34.
14. Parak J, Korhonen I. Evaluation of wearable consumer heart rate monitors based on photopletysmography. *Conf Proc IEEE Eng Med Biol Soc*. 2014;2014:3670–3.
15. Parak J, Tarniceriu A, Renevey P, Bertschi M, Delgado-Gonzalo R, Korhonen I. Evaluation of the beat-to-beat detection accuracy of PulseOn wearable optical heart rate monitor. *Conf Proc IEEE Eng Med Biol Soc*. 2015;2015:8099–8102.
16. Rainmaker DC. *Polar A360 In-Depth Review*. [cited 1 Mar 2017]. Available from: http://www.dcrainmaker.com/2015/12/polar-a360-depth-review.html.
17. Schäfer A, Vagedes J. How accurate is pulse rate variability as an estimate of heart rate variability? A review on studies comparing photoplethysmographic technology with an electrocardiogram. *Int J Cardiol*. 2013;166:15–29.
18. Spierer DK, Rosen Z, Litman LL, Fujii K. Validation of photoplethysmography as a method to detect heart rate during rest and exercise. *J Med Eng Technol*. 2015;39(5):264–71.
19. Stahl SE, An HS, Dinkel DM, Noble JM, Lee JM. How accurate are the wrist-based heart rate monitors during walking and running activities? Are they accurate enough? *BMJ Open Sport Exerc Med*. 2016;2(1):e000106.
20. Stern J. Fitness bands with heart-rate tracking are missing a beat. *The Wall Street Journal*. [cited 1 Mar 2017]. Available from: http://www.wsj.com/articles/fitness-bands-with-heart-rate-tracking-are-missing-a-beat-1418759929.
21. Tuulari E. Polar wrist heart rate. *Polar Electro White Paper*. 2015 June 5.
22. Vander AJ, Sherman JH, Luciano DS. *Human Physiology*. 3d ed. New York: McGraw-Hill; 1980. p. 283.
23. Wang R, Blackburn G, Desai M, et al. Accuracy of wrist-worn heart rate monitors. *JAMA Cardiol*. 2017;2(1):104–6.