Janicki et al. (2020) *Int J Res Ex Phys*. 15(2):91-105.

Sponsored by:
Exercise and Sport Science Program
Western State Colorado University

International Journal of Research in Exercise Physiology

Original Research Article

The Accuracy of Heart Rate and Energy Expenditure Measurements using the PolarM430 Watch during Exercise

Garrett J. Janicki[1], John P. Porcari[1], Andrew R. Jagim[1-2], Kari Emineth[1], Scott Dorberstein[1], Elisabeth Schwab[1], Carol Foster[1]

[1]Department of Exercise and Sport Science, University of Wisconsin-La Crosse, La Crosse, WI, USA

[2]Sports Medicine, Mayo Clinic Health System, La Crosse, WI, USA

Abstract

Purpose: The aim of the current study was to assess the accuracy of the Polar M430 fitness watch for the measurement of heart rate (HR) and energy expenditure (EE) during low, moderate, and high intensity treadmill exercise. **Methods:** Twenty subjects (males, n = 12; Age: 24.2 ± 2.41 yrs., Height: 175.7 ± 6.85 cm, Weight: 88.1 ± 13.02 kg; BMI: 28.44 ± 2.90 kg/m^2) (females, n = 8; Age: 22.5 ± 2.98 yrs., Height: 171.9 ± 4.48 cm, Weight: 73.3 ± 9.10 kg; BMI: 24.81 ± 3.00 kg/m^2) completed a single bout of exercise on a treadmill at three incremental intensities using self-selected ratings of perceived exertion (RPE) - derived workloads at RPE 11 (low intensity), RPE 13 (moderate intensity), and RPE 15 (high intensity), respectively. Each workload was 5-minutes in duration and subjects wore a PolarM430 watch throughout the entire protocol. An electrocardiogram (ECG) and indirect calorimetry were used as criterion measures of HR and EE, respectively. Two-way ANOVA with repeated measures was used to examine differences in HR and EE between the PolarM430 and criterion measures. Pairwise comparisons were made using Tukey's post-hoc tests. Intraclass correlation coefficients were used to examine the relationship between HR and EE. Data are presented as mean ± 95% confidence intervals. **Results:** The PolarM430 significantly underestimated HR at RPE 15 -3.20 (-5.00, -1.40) bpm. ICC analysis indicated an excellent level of agreement between the PolarM430 and criterion measures for HR (ICC =.96-.99). The PolarM430 significantly underestimated EE by -19.25 (-14.77, -3.37) kcals (14.9%) at RPE 15 and -10.5 (-20.09, -0.91) kcals (-8.78%) throughout the entire testing protocol. ICC analysis indicated a moderate to good level of agreement (ICC = .66-.89) for EE. **Conclusions:** Based on these results, it appears the Polar M430 fitness watch is an accurate tool to measure HR at low to moderate exercise intensities, but may underestimate HR at higher exercise intensities. The PolarM430 watch appears to be an accurate tool to measure EE during low and moderate intensity exercise, but may underestimate EE at higher exercise intensities, as well as over the course of an entire bout of exercise.

Key Words: Biofeedback, Fitness Watch, Wearable Technology.

Introduction

Wearable technology continues to be one of the top fitness trends and is estimated to be a $95 billion industry[1]. Wearable technology, which includes fitness trackers, smart watches, and fitness watches etc., often provides the consumer with various biofeedback metrics such as heart rate (HR), energy expenditure (EE) and sleep. Certain devices also contain built in global position systems capabilities to provide specific movement characteristic (i.e. speed, distance, pace, etc.) profiles during outdoor activities. Some of the most popular wearable technology devices are fitness watches, which offer the basic functions of a watch in addition to the biofeedback parameters that often appeal to fitness enthusiasts and athletes. As the technology and software of these devices becomes more advanced, it is important to continue examining the accuracy of these devices to ensure the biofeedback derived from the devices are an accurate representation of the consumer's physiological responses, which are starting to become integrated into health monitoring, research, sport programming and weight loss interventions.

The majority of fitness watches rely on photoplethysmography (PPG) for the determination of heart rate (HR), which utilizes light refraction from the capillaries to sense blood volume changes which in turn can be used to estimate heart rate[2]. Photoplethysmography allows for users to measure their HR with a fitness watch or similar wrist worn device, while allowing for ambulatory movement and more versatility during activities compared to a chest-strap or other direct measure of HR. However, the accuracy of PPG-based measures of HR may be influenced by a variety of factors such as motion artifact, hydration status, blood flow, skin color and fit of device[2]. Therefore it is important to examine the accuracy of wearable devices which utilize PPG-derived measures of HR under a variety of conditions and across different populations, as each setting may alter the validity the device. Previous studies have examined the accuracy of wrist-worn devices when measuring HR during exercise with conflicting findings. Recently Benedetto et al.[3] found a mean bias of -6 bpm; however, the authors noted poor precision of measurement with lower and upper limits of agreement ranging from -28.5 to 16.8 bpm when a wrist-worn activity tracker was compared to criterion measures during a 10-minute bout of exercise on a cycle ergometer at a self-selected intensity. These results indicate that wrist-worn devices may significantly underestimate HR during a wide range of exercise intensities, which may create problems for individuals relying on immediate biofeedback of HR measures for intensity-guided workouts. Conversely, Scherbina et al.[4] observed an error of 1.8% in HR during an incremental bout of cycling compared to 5.5% during walking. However all wrist-worn devices assessed in the study by Shcherbina et al.[4] yielded HR values with acceptable error ranges (<5% determined a priori) under controlled laboratory conditions, thereby indicating exercise modality and testing protocol likely influence the accuracy of each device. However, it is up to the consumer's discretion regarding what is an acceptable range or error when classifying a device as accurate or not, which is likely dependent upon the user's goal and how the results are being interpreted and used in HR guided exercise programming. Furthermore, it does not appear as though wrist-worn devices consistently underestimate HR during exercise, as Thiebaud et al.[5] found that two out of three wrist-worn devices actually overestimated HR with a mean absolute percentage error ranging from 1-8% during an incremental treadmill test. This inconsistency creates problems for how these devices are used in various settings, as poor reliability from device to device makes it difficult to extrapolate biofeedback across multiple devices, particularly when it is unknown whether a measured HR could be higher or lower than actual HR. These inconsistencies warrant further research to examine the accuracy of newer wrist-worn devices to better our understanding of how exercise intensity influences the accuracy of HR measurements.

In addition to HR, several wrist-worn devices also provide estimates of EE, both continuously throughout the day and during a distinct bout of physical activity. Previous research has observed significant discrepancies between criterion measures of EE and wrist-worn devices during exercise under controlled laboratory settings and free-living daily activities.[5-7] Wrist-worn devices rely on proprietary algorithms to estimate EE using a combination of demographic, accelerometry and HR-derived information from the user. Therefore, the accuracy of each device might provide variable estimates of EE, depending on the specific formulas used and how each variable is weighted. For example, Thiebaud et al.[5] observed a wide range of error across three wrist-worn devices ranging from 1-9% for another device and up to 25-61% for another during an incremental exercise protocol on a treadmill. Wallen et al.[7] noted similar findings in subjects who wore four different wrist-worn monitors at rest and during exercise with the percentage error for EE estimation ranging from 9 - 43% across the different devices. In one of the largest and most well-controlled studies to date, Murakami et al.[8] observed a high degree of variability in EE across 12 commercially available fitness trackers during a standardized day in a metabolic chamber and during a 15-day ambulatory setting using doubly labeled water. Specifically, the authors[8] reported that the physical activity EE values from 9 out of 12 devices tested were significantly different than the metabolic chamber physical activity EE and none of the devices correlated well to the physical activity EE from the chamber (r = .13 - .37). Additionally, the authors found that 10 out of the 12 devices significantly underestimated EE during the 15-day free living period compared to doubly labeled water. With such high variability from device to device and an overall low degree of accuracy compared to criterion measures of EE, relying on wrist-worn devices for accurate estimates of EE may be problematic and consumers may want to exercise caution regarding how these data are used in practice. However, each device may provide varying degrees of accuracy and therefore it is important to further examine specific devices under controlled laboratory conditions to determine the validity of each device during different exercise intensities and modalities.

A popular wrist-worn device on the market marketed towards active individuals and athletes is the PolarM430 (*Polar Electro Oy, Kempele, Finland*). As with any device, the accuracy of this product should be examined to understand the utility of the device in practice. Therefore, the purpose of this study was to assess the accuracy of the PolarM430 fitness watch for the assessment of HR and EE during low, moderate, and high intensity treadmill exercise.

Methods

Experimental Design

Subjects reported to the Human Performance Laboratory to complete a familiarization session. During familiarization, subjects were instructed on how to use the Borg 6-20 rating of perceived exertion (RPE) scale[9] and to determine the appropriate treadmill speed/grade combinations to elicit exercise intensities that corresponded to RPE values of 11 (light), 13 (moderate), and 15 (hard), which would be used for future testing. Subjects returned to the laboratory within 7 days for experimental testing. On the day of testing, subjects were asked to report to the laboratory in a post-absorptive state (> 2 hours) and to have refrained from strenuous exercise for the previous 24 hours. Subjects were equipped with the PolarM430 watch and criterion devices for determination of HR and EE, followed by the completion of a 15-minute treadmill protocol at incremental exercise intensities. Heart rate and EE expenditure were recorded at the end of each five-minute stage.

Subjects

Twenty apparently healthy and recreationally active men and women between 18 - 35 yrs. of age were recruited to participate in the study. Descriptive characteristics of the subjects are presented in Table 1.

Table 1. Descriptive characteristics of subjects (N = 20).

	Males (n = 12)	Females (n = 8)
Age (yrs)	24.2 ± 2.41	22.5 ± 2.98
Height (cm)	175.7 ± 6.85	171.9 ± 4.48
Weight (kg)	88.1 ± 13.0	73.3 ± 9.1
BMI (kg/m^2)	28.4 ± 2.90	24.8 ± 3.00

Values represent mean ± SD.

Inclusion criteria included being between the ages of 18 - 35 years, recreationally active and free of contraindications to participation in exercise. Recreationally active was defined as engaging in regular aerobic training at least three days per week during the previous six months. Potential subjects completed a Par-Q and health history questionnaire to screen for cardiovascular and orthopedic contraindications to exercise. The study protocol was approved by the University Institutional Review Board for the Protection of Human Subjects. Eligible participants were provided with a detailed description of the testing protocol and expectations for study participation prior to providing written informed consent. During the familiarization session, basic demographic information was collected and later used to set up a user profile within the data acquisition software program linked with the PolarM430 watch.

Criterion Measures

An electrocardiogram (ECG) using a modified 3 limb lead system (*CASE system, General Electric, Boston, MA*) served as the criterion measure of HR. Indirect calorimetry was used as the criterion measure of EE (*ParvoMedics, Sandy, UT*). Prior to each testing session, the metabolic cart was calibrated with gases of known concentrations (15.98% O_2 and 4.12% CO_2), as well as with room air (20.93% O_2 and 0.03% CO_2), as per manufacture guidelines. Calibration of the pneumotachometer was completed using a 3 Liter calibration syringe.

Procedures

Height was assessed using a wall mounted stadiometer and weight was measured using a Rice Lake Digital Scale (*Rice Lake Weighing Systems, Rice Lake, WI*). Subjects completed a 5-minute warm-up on the treadmill which consisted of brisk walking. After the warm-up, subjects were equipped with the PolarM430 watch (*Polar Electro Oy, Kempele, Finland*) on their left wrist according to manufacturer recommendations. Subjects then completed three, 5-minute stages on a motorized treadmill at incremental workloads corresponding to RPE values of 11 (light), 13 (moderate), and 15 (hard). During the treadmill exercise, HR was measured continuously using both ECG and the PolarM430. Simultaneous HR values from the ECG recording and the PolarM430 at the end of the 4th and 5th minute of each stage were averaged and presented as the mean HR for each exercise intensity. Total accumulated EE was recorded upon completion of each 5-minute stage from both the metabolic cart and the PolarM430. For the metabolic cart, EE was calculated using indirect calorimetry-derived respiratory exchange ratio values according to previous methods[10]. At the end of the third stage, subjects walked for 5 minutes at a self-selected pace as a cool down.

Statistical analyses

Standard descriptive statistics were used to characterize the subject population and summarize the HR and EE data. Initially, a three-way repeated measures ANOVA with repeated measures was used to compare directly measured and watch measured HR and EE between males and females across RPE levels. Data between the watch and directly measured HR and EE across RPE levels were then compared using two-way ANOVA with repeated measures. Significant F ratios were followed by pairwise comparisons using Tukey's post-hoc tests. Intraclass correlation coefficients were also used to examine the relationship between directly measured and watch HR and EE. Consistent with previous literature[11], values less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.90 were classified as poor, moderate, good, and excellent reliability, respectively. Mean bias was calculated as actual (ECG or indirect calorimetry) – predicted and data are presented as mean ± 95% confidence intervals (CI). Bland-Altman plots were used to assess the levels of agreement between the HR and EE values for the PolarM430 and criterion measures. Alpha was set at $p < 0.05$ to achieve statistical significance. Data analysis was conducted using the Statistical Package for the Social Sciences (*IBM SPSS Statistics for Windows, Version 25.0; IBM Corp., Armonk, NY*).

Results

The three-way repeated measures ANOVA indicated no significant difference between males and females, therefore, data were collapsed across gender for subsequent analysis.

Heart rate

HR values for the PolarM430 and criterion measures across each intensity are presented in Table 2. A significant difference in HR between the PolarM430 and criterion measures was observed with the watch underestimating HR by -3.2 (95% CI: -5.00, -1.40) bpm at RPE 15, and -2.55 (95% CI: -3.56, -1.54) throughout the entire session. ICC analysis indicated an excellent level of agreement at each exercise intensity with ranges of ICC =.96 -.99 (Table 2). Figure 1 presents the relationship between heart rate vales from the Polar M430 and ECG.

The Bland Altman plot in Figure 2 indicated a non-significant relationship for mean bias versus the ECG criterion method for the PolarM430 watch ($r = -0.202$, $p = 0.12$) with levels of agreement ranging from -11.12 to 6.02 bpm.

Table 2. Comparison of HR between the Polar M430 watch and criterion measure.

	PolarM430	ECG	Mean Bias	95% CI	ICC
Hear Rate (bpm)					
RPE 11	108.5 ± 15.35	110.2 ± 16.44	-1.70	-3.79, 0.39	.96
RPE 13	133.8 ± 22.69	136.5 ± 22.23	-2.75	-5.01, -0.49	.98
RPE 15	151.6 ± 22.11*	154.8 ± 23.52	-3.20	-5.00, -1.40	.99
Entire Test (Avg)	131.3 ± 26.77*	133.8 ± 27.65	-2.55	-3.56, -1.54	.99

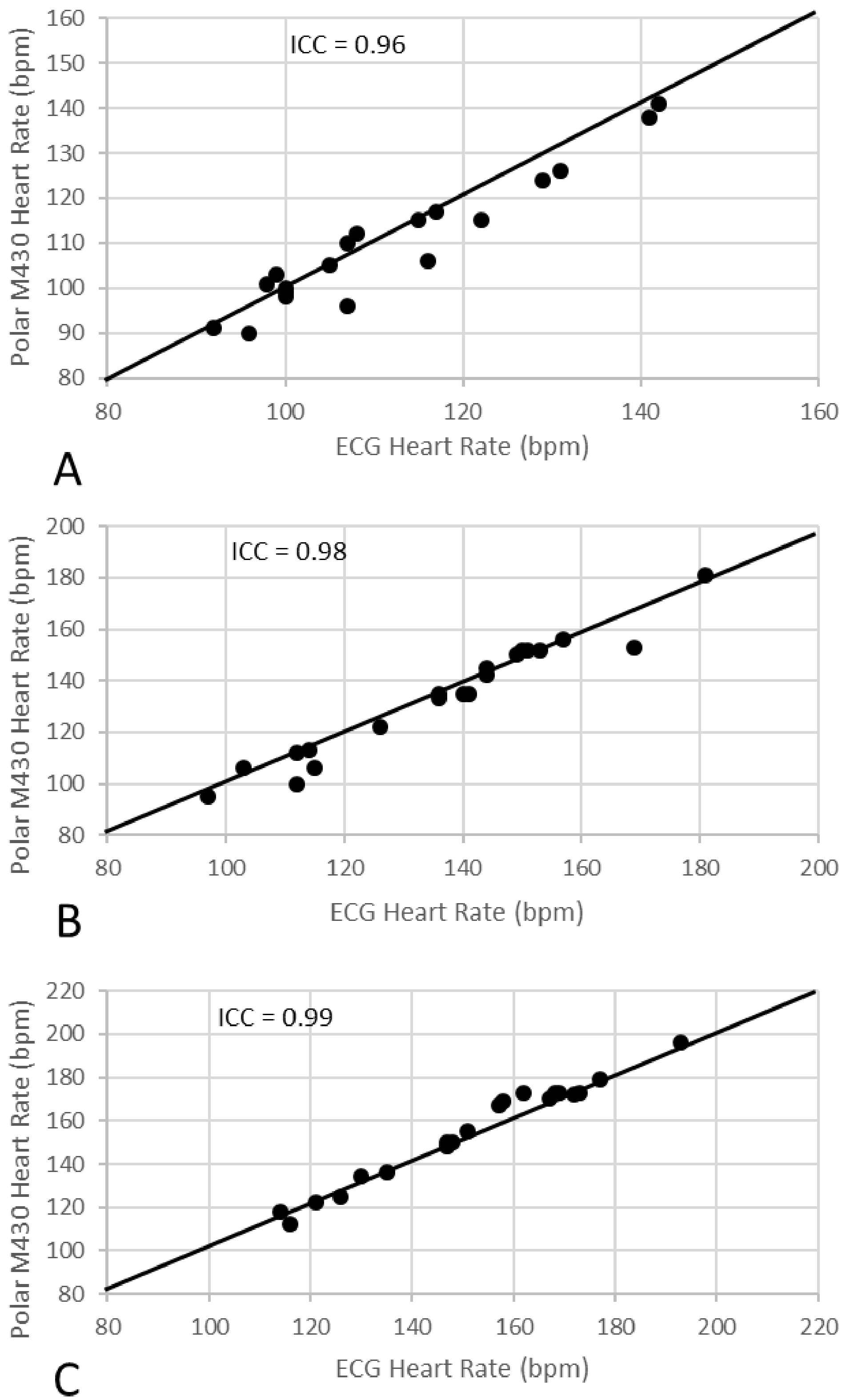

Figure 1. Heart rate values for PolarM430 and ECG (bpm) for the RPE 11 (A), 13 (B) and 15 (C) stages.

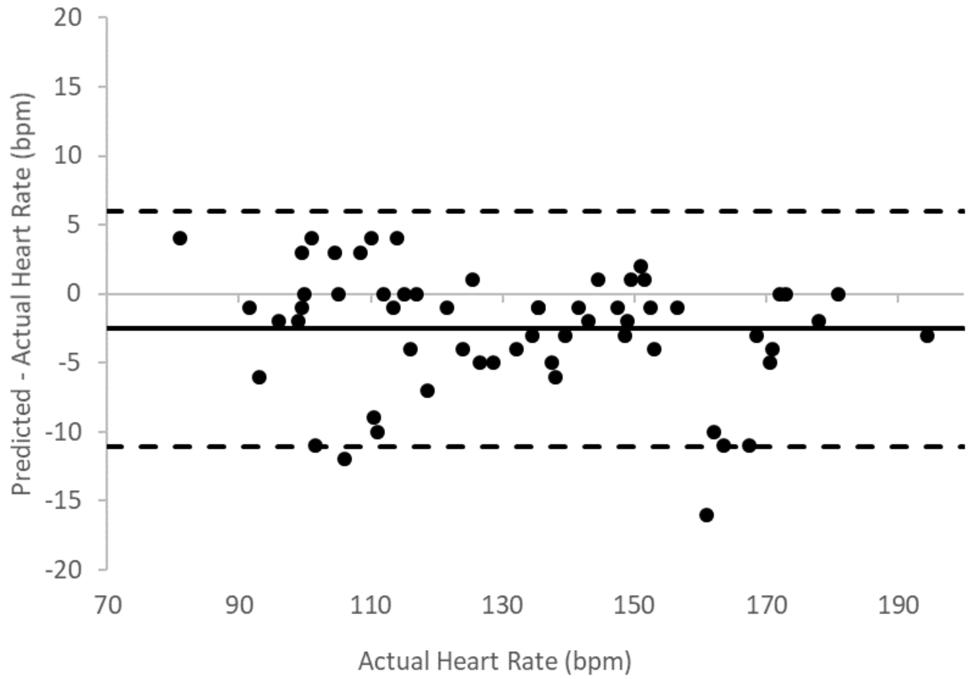

Figure 2. Relationships between constant error (actual – predicted HR) and average HR ([measured + predicted]/2). Solid line represents mean bias. Dashed lines represent ± 1.96 SD of mean bias. HR = heart rate.

Energy Expenditure

EE values for the PolarM430 and criterion measures across each exercise intensity are presented in Table 3. The PolarM430 significantly underestimated EE by -9.25 (95% CI: -14.77, -3.73) kcals (-18.3%) at RPE 15 and -10.5 (-20.09, -0.91) kcals (-8.78%) throughout the entire protocol. ICC analysis indicated a good level of agreement at RPE 13 (ICC = .89), a moderate relationship was observed at RPE 11 (ICC =.66) and 15 (ICC = .78), respectively. Figure 3 presents the relationship between energy expenditure values from the Polar M430 and indirect calorimetry.

The Bland Altman plot in Figure 4 indicated a significant relationship for mean bias versus the indirect calorimetry criterion method for the PolarM430 watch ($r = -0.52$, $p < 0.01$) with levels of agreement ranging from -22.57 to 15.58 kcals.

Table 3. Comparison of EE between the Polar M430 watch and criterion measure.

	PolarM430	Indirect Calorimetry	Mean Bias	95% CI	ICC
Energy Expenditure (kcals)					
RPE 11	31.2 ± 8.88	29.7 ± 8.90	1.55	-1.87, 4.97	.66
RPE 13	43.7 ± 12.35	46.5 ± 13.95	-2.80	-5.76, 0.16	.89
RPE 15	52.5 ± 13.5*	61.7 ± 18.73	-9.25	-14.77, -3.73	.78
Entire Test (Sum)	127.3 ± 33.56*	137.8 ± 39.05	-10.50	-20.09, -0.91	.85

Values represent mean ± SD. 95% CI = 95% Confidence Interval. ICC = Intraclass correlation coefficient.
* Significant difference from criterion measure (p<.05).

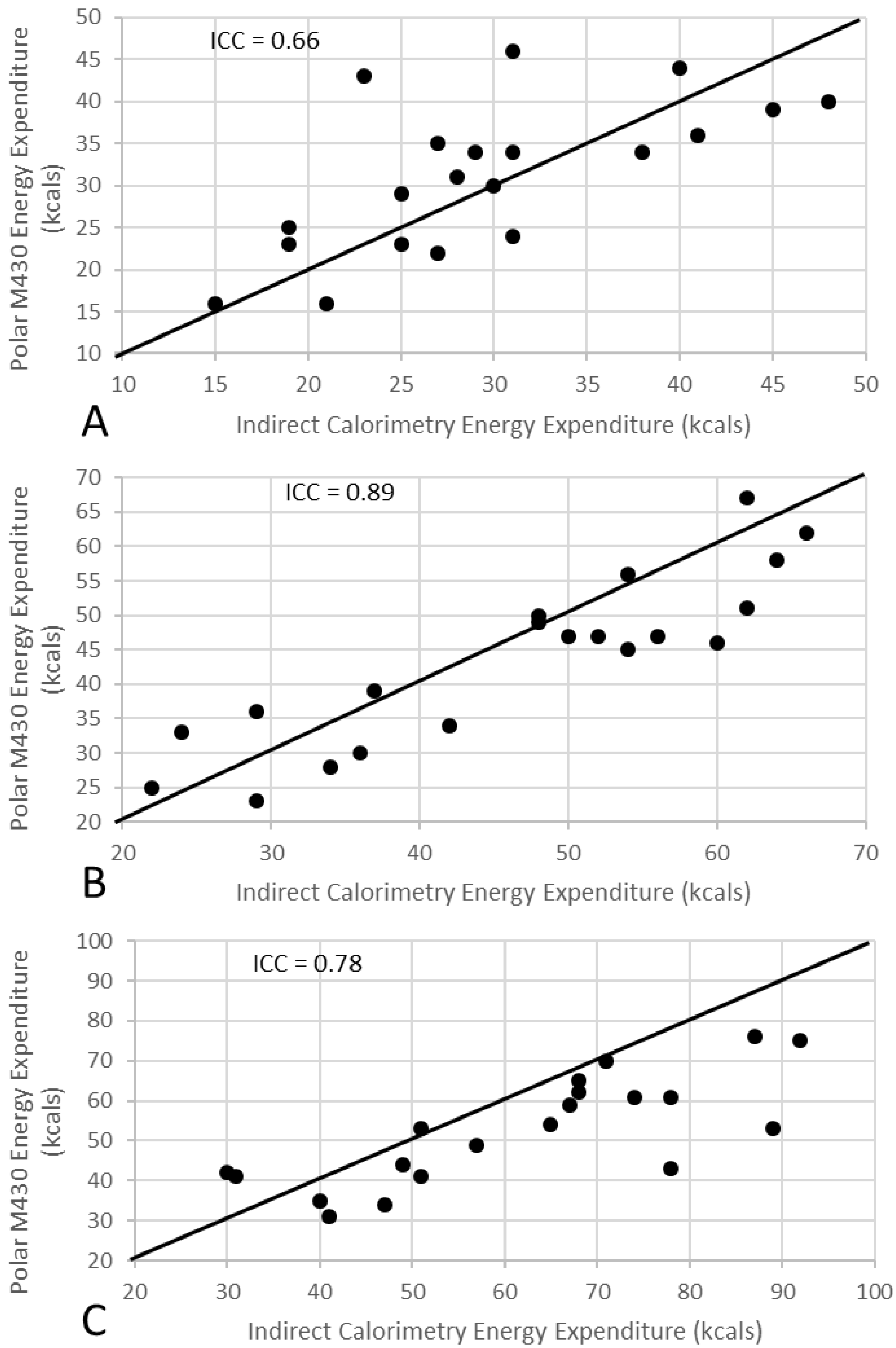

Figure 3. Energy Expenditure values for PolarM430 and indirect calorimetry (kcals) for the RPE 11 (A), 13 (B) and 15 (C) stages.

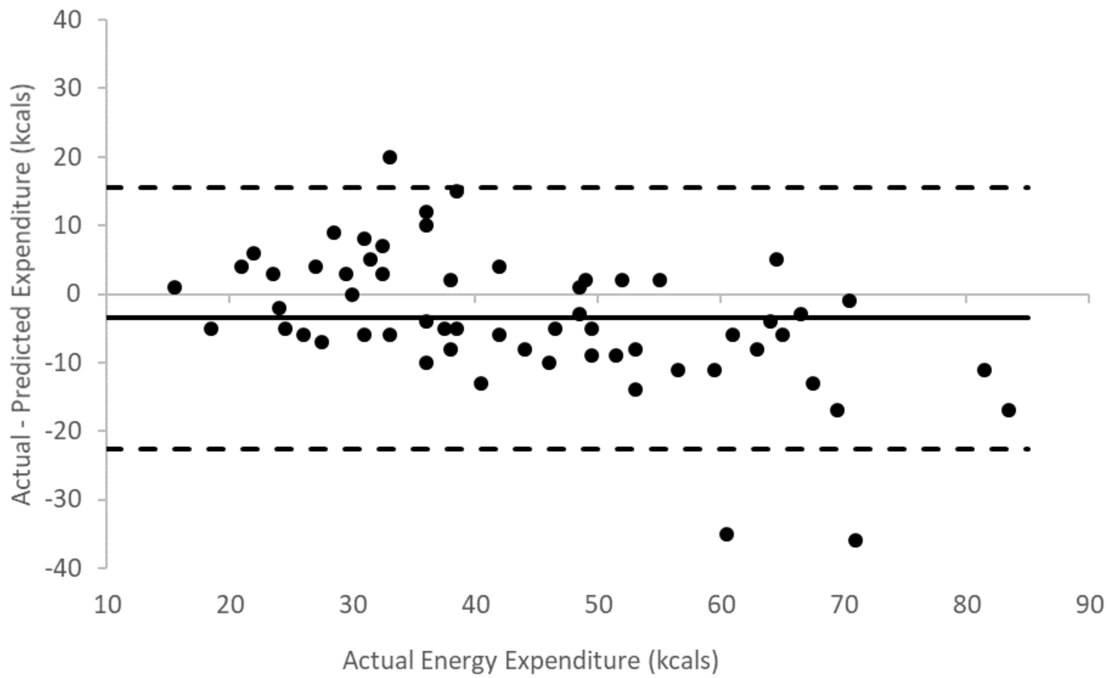

Figure 4. Relationships between constant error (actual – predicted EE) and average EE ([measured + predicted]/2). Solid line represents mean bias. Dashed lines represent ± 1.96 SD of mean bias. EE = energy expenditure.

Discussion

The primary aim of the current study was to assess the accuracy of the PolarM430 watch for the measurement of HR and EE across a range of exercise intensities. The main findings from the current study indicate the PolarM430 watch provided an accurate measure of HR and EE at low and moderate intensity exercise, but became less accurate during high intensity exercise for both measures. The PolarM430 significantly underestimated HR at RPE 15 (high intensity); however, only a modest mean bias of -3 bpm (-2.0%) was observed. Furthermore, the watch yielded limits of agreement ranging from -11.12 (-9.1%) to 6.02 (5.2%) bpm, which may be considered an acceptable degree of error depending on user preference and how the data are being utilized.

The results of the current study indicate lower levels of mean bias and narrower limits of agreement for HR analysis compared to other devices that have previously been examined within the literature[3,12]. For example, Benedetto et al.[3] reported a mean bias of 6 bpm and limits of agreement of -28.5 to 16.8 bpm when a Fitbit device was compared to criterion measures of ECG during a 10-minute bout of exercise at self-guided intensities on a cycle ergometer. Similarly, Cadmus-Bertram et al.[12] reported limits of agreement ranging from -17 to 22 bpm across four different wrist-worn activity trackers when subjects were at rest, with values ranging from -41 to 39 bpm when subjects exercised at 65% of max heart rate. Therefore, it appears the PolarM430 watch may provide more accurate measures of HR compared to other commercially available wrist-worn activity trackers.

In the current study, as exercise intensity increased, the PolarM430 provided less accurate measures of HR. Specifically, the PolarM430 watch tended to be the most accurate at lower intensities as the watch exhibited a mean bias of -1.70 bpm and -2.75 bpm for low and moderate intensity exercise, respectively, with a higher mean bias of -3.2 bpm observed during the high exercise intensity stage. These findings are in alignment to a previous study, which also demonstrated decreases in accuracy during higher exercise intensities[13]. Jo et al.[13] similarly reported a mean bias of -5.3 bpm (8.5%) and limits of agreement of -26 to 16 bpm for a Fitbit device when subjects were exercising at an ECG-determined HR < 117 bpm and a higher mean bias of -13 bpm (9.8%) and wider limits of agreement of -54 to 29 bpm when exercising at ECG determined HR > 116 bpm. Although the results by Jo et al.[13] follow a similar pattern to those observed in the current study, the mean bias and levels of agreement from the current study indicate a much more favorable level of accuracy for the PolarM430 compared to the Fitbit. In contrast, Stahl et al.[14] observed lower error values at higher intensities, compared to those observed at lower intensities, in subjects completing a graded exercise treadmill test using a variety of wrist-worn activity trackers compared to criterion measures. Specifically, Stahl et al.[14] reported mean absolute percentage errors (MAPE) values of 5-8% (one device was an outlier with MAPE values of 16%) at 3.2 km/hr compared to 1-3% at 9.6 km/hr. It is possible that at higher exercise intensities a higher cardiac output enhances the ability of the LED sensors to detect the backscatter of light and create a more favorable profile for PPG-derived measures of HR for certain devices. Interestingly, Thiebaud et al.[5] observed variable mean biases and MAPE values across a graded exercise treadmill test in multiple wrist worn devices. Specifically, the authors [5] noted the highest mean bias of 7 bpm and MAPE of 7-8% at 4.8 km/hr compared to a mean bias of 1-5 bpm and MAPE of 1-6% at 3.2 km/hr, and a mean bias of -5 - 2 bpm and MAPE of -3.5 - 1% at 9.7 km/hr, indicating inconsistent patterns of accuracy at different exercise intensities. This type of pattern makes it difficult to account for any measurement errors using any kind of a correction factor. The variable degrees of accuracy from each device may be attributable to several factors that may influence the ability of PPG to accurately measure HR. For example, each specific device may utilize different light sensors or fit differently on a subject's arm, thereby hindering the ability to accurately measure HR. Other factors such as hydration status, motion artifact and skin color have all been shown to influence the accuracy of PPG-derived measures of HR under various conditions; all of which may have been confounding factors influencing the accuracy of HR assessment in the current study as well[2,15]. Furthermore, despite the worsening mean bias observed in the current study as exercise intensity exercised (Figure 2), on an individual level there are cases where there may be better or worse agreement. Therefore it is possibly that wearable technologies may be more or less accurate and functional for certain individuals, depending on the presence of the aforementioned confounding variables, intensity and modality of exercise being performed, and how the data is being utilized. For example, a larger individual with higher levels of body fatness and darker skin complexion may not experience as adequate levels of HR accuracy from wrist-worn devices compared to a leaner individual with lighter skin complexion. Additionally, someone using wrist-worn devices for HR-guided intensity selection during high intensity interval training may be misled with underestimated HR readings; all of which may be exacerbated with exercised modalities requiring a high degree of arm movement or hydration status[2]. Future research is needed to identify the usability of various wearable technologies across diverse populations and under varying exercise conditions.

Another aim of the current study was to examine the accuracy of the PolarM430 watch for the estimation of EE during exercise. The results of the current study indicate the PolarM430 was able to accurately measure EE at low and moderate intensity exercise, but failed to do so during high intensity exercise. Specifically, the PolarM430 watch became less accurate as the exercise intensity increased, as evidence by the error percentage values of 5.1%, 6.0%, and 14.9%, for low, moderate, and high intensity exercise, respectively. Overall, the error percentages observed in the current study are below those previously reported when fitness watches or wrist-worn activity trackers were compared to criterion measures during exercise. Previous studies have found percentage error ranges of 16-60% depending on the device and intensity of exercise[4,5,7,16]. Shcherbina et al.[4] reported median error rates of 27% for a Fitbit Surge and 92% for a PulseOn device during cycling and walking, both of which exceed the a priori determined threshold of 10% established by the authors for an acceptable error range. Thiebaud et al.[5] also noted a wide range in the accuracy of EE estimation for multiple wrist worn activity trackers. Specifically, the authors[5] observed MAPE values of 44%, 0.4% and -9.4% at 3.2 km/hr for a Fitbit Surge, TomTom Cardio and Microsoft Band device, respectively. These devices continued to exhibit wide ranges of error at higher exercise intensities with MAPE values of 37%, -16.4% and 4% at 8 km/hr for the Fitbit Surge, TomTom Cardio and Microsoft Band devices, respectively; thereby, indicating that wrist-worn devices may over or underestimate EE during exercise. Wallen et al.[7] also reported low levels of accuracy from multiple wrist-worn devices when estimating EE during multiple modes of activity with limits of agreement ranging from -267 to 66 kcals[7]. As a lot of activity trackers and fitness watches may be used to estimate EE during specific activities or throughout the day to better direct dietary practices, these observed inaccuracies in EE measurements could be problematic. The majority of devices appear to overestimate EE, which may provide consumers with an inflated sense of total daily activity energy expenditure, which could subsequently contribute to excess energy intake.

It is difficult to specifically determine the source of error regarding the estimation of EE as each devise relies on proprietary algorithms to calculate EE. Most algorithms likely include several fixed factors such as age, height and weight as part of the equation, however it is unknown if each device also integrates HR or make adjustments based on the mode of activity being performed. In the current study, HR measurements were less accurate at the higher exercise intensity and tended to underestimate HR, which may have contributed to the underestimation in EE at the same intensity; that is, assuming HR is included in the algorithm and how it is weighted in the equation. It is also possible that EE values are impacted by accelerometer-derived information which could also influence the accuracy of each device depending on the mode of activity. For example, in the current study, most subjects completed the 11 and 13 RPE stages at similar speed and grades, respectively; however, at the 15 RPE stage, subjects utilized varying speeds and grades to elicit the necessary perceived intensity. As a result, some subjects chose to walk at high inclines to achieve a 15 RPE while others chose to run at low inclines. Therefore, it is possible that the variance in speed and grade usage among subjects for this stage may have influenced the accuracy of EE estimation. Previous research has indicated that devices which rely on accelerometry to estimate EE have been shown to have varying levels of accuracy across different modes of activity, due to varying degrees of bodily movement[4,17,18]. Future research should examine the accuracy of fitness watches across multiple exercise modalities and intensities of effort.

A limitation of the current study was that the subject population consisted of young, healthy and predominantly Caucasian adults. Therefore the results of the current study may not translate across other populations, particularly those of darker skin color as the PPG technology is reliant on light emission which has been shown to be influenced by skin tone[2]. Additionally, only a single mode of activity was performed in the current study, which is another limitation as the accuracy of the device may vary with different exercise modalities. Further, certain devices may include a specific setting for each specific mode of exercise, while others may not, thereby invalidating the applicability of the proprietary algorithm for that activity as previously mentioned[5]. The use of a graded exercise protocol instead of intervals is another limitation. It is currently unknown whether fitness watches and other wrist worn devices are able to rapidly respond to abrupt changes in exercise intensities which may limit their use during HR-guided exercise protocols such as interval training.

Conclusions

In summary, results of the current study indicate that the Polar M430 fitness watch is able to provide an accurate measure of HR and EE during low and moderate intensity exercise on a treadmill. This device can help active individuals to monitor HR during bouts of exercise to determine exercise intensity and characterize training sessions over time. Further, estimations of EE from the device can be used to direct weight management efforts or target dietary strategies to account for energy expenditure at rest and low to moderate intensity bouts of exercise. However, caution should be taken when using this device to monitor HR and EE at higher exercise intensities, as levels of accuracy may diminish as exercise intensity increases.

Disclosures

No funding was provided in support of this study. The authors have no conflicts of interest to disclose.

Address for Correspondence

Andrew Jagim, Ph.D., Director of Sports Medicine Research, Mayo Clinic Health System, 700 West Ave. S., La Crosse, WI, United States, 54601. Phone: 608-392-5280; Email: Jagim.andrew@mayo.edu.

References

1. Thompson W. Worldwide survey of fitness trends for 2020. *ACSM's Health & Fitness Journal.* 2019;23(6):10-18.
2. Castaneda D, Esparza A, Ghamari M, Soltanpur C, Nazeran H. A review on wearable photoplethysmography sensors and their potential future applications in health care. *Int J Biosens Bioelectron.* 2018;4(4):195-202.
3. Benedetto S, Caldato C, Bazzan E, Greenwood DC, Pensabene V, Actis P. Assessment of the Fitbit Charge 2 for monitoring heart rate. *PLoS One.* 2018;13(2):e0192691.
4. Shcherbina A, Mattsson CM, Waggott D, et al. Accuracy in Wrist-Worn, Sensor-Based Measurements of Heart Rate and Energy Expenditure in a Diverse Cohort. *J Pers Med.* 2017;7(2).
5. Thiebaud RS, Funk MD, Patton JC, et al. Validity of wrist-worn consumer products to measure heart rate and energy expenditure. *Digit Health.* 2018;4:2055207618770322.
6. Murakami H, Kawakami R, Nakae S, et al. Accuracy of 12 Wearable Devices for Estimating Physical Activity Energy Expenditure Using a Metabolic Chamber and the Doubly Labeled Water Method: Validation Study. *JMIR Mhealth Uhealth.* 2019;7(8):e13938.
7. Wallen MP, Gomersall SR, Keating SE, Wisloff U, Coombes JS. Accuracy of Heart Rate Watches: Implications for Weight Management. *PLoS One.* 2016;11(5):e0154420.
8. Murakami H, Kawakami R, Nakae S, et al. Accuracy of Wearable Devices for Estimating Total Energy Expenditure: Comparison With Metabolic Chamber and Doubly Labeled Water Method. *JAMA Intern Med.* 2016;176(5):702-703.
9. Borg GA. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc.* 1982;14(5):377-381.
10. Weir JB. New methods for calculating metabolic rate with special reference to protein metabolism. *J Physiol.* 1949;109(1-2):1-9.
11. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med.* 2016;15(2):155-163.
12. Cadmus-Bertram L, Gangnon R, Wirkus EJ, Thraen-Borowski KM, Gorzelitz-Liebhauser J. Accuracy of Heart Rate Monitoring by Some Wrist-Worn Activity Trackers. *Ann Intern Med.* 2017;167(8):607-608.
13. Jo E, Lewis K, Directo D, Kim MJ, Dolezal BA. Validation of Biofeedback Wearables for Photoplethysmographic Heart Rate Tracking. *J Sports Sci Med.* 2016;15(3):540-547.

14. 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.
15. 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):264-271.
16. Nelson MB, Kaminsky LA, Dickin DC, Montoye AH. Validity of Consumer-Based Physical Activity Monitors for Specific Activity Types. *Med Sci Sports Exerc.* 2016;48(8):1619-1628.
17. Swan PD, Byrnes WC, Haymes EM. Energy expenditure estimates of the Caltrac accelerometer for running, race walking, and stepping. *Br J Sports Med.* 1997;31(3):235-239.
18. Stackpool C, Porcari, JP., Mikat, RP., Gillete, C., ,Foster, C. The accuracy of various activity trackers in estimating steps taken and energy expenditure. *J Fit Research.* 2014;3(3):32-48.