

Validity of Wearable Activity Monitors during Cycling and Resistance Exercise

BENJAMIN D. BOUDREAUX, EDWARD P. HEBERT, DANIEL B. HOLLANDER, BRIAN M. WILLIAMS, CORINNE L. CORMIER, MILDRED R. NAQUIN, WYNN W. GILLAN, EMILY E. GUSEW, and ROBERT R. KRAEMER

Department of Kinesiology and Health Studies, Southeastern Louisiana University, Hammond, LA

ABSTRACT

BOUDREAUX, B. D., E. P. HEBERT, D. B. HOLLANDER, B. M. WILLIAMS, C. L. CORMIER, M. R. NAQUIN, W. W. GILLAN, E. E. GUSEW, and R. R. KRAEMER. Validity of Wearable Activity Monitors during Cycling and Resistance Exercise. *Med. Sci. Sports Exerc.*, Vol. 50, No. 3, pp. 624–633, 2018. **Introduction:** The use of wearable activity monitors has seen rapid growth; however, the mode and intensity of exercise could affect the validity of heart rate (HR) and caloric (energy) expenditure (EE) readings. There is a lack of data regarding the validity of wearable activity monitors during graded cycling regimen and a standard resistance exercise. The present study determined the validity of eight monitors for HR compared with an ECG and seven monitors for EE compared with a metabolic analyzer during graded cycling and resistance exercise. **Methods:** Fifty subjects (28 women, 22 men) completed separate trials of graded cycling and three sets of four resistance exercises at a 10-repetition-maximum load. Monitors included the following: Apple Watch Series 2, Fitbit Blaze, Fitbit Charge 2, Polar H7, Polar A360, Garmin Vivosmart HR, TomTom Touch, and Bose SoundSport Pulse (BSP) headphones. HR was recorded after each cycling intensity and after each resistance exercise set. EE was recorded after both protocols. Validity was established as having a mean absolute percent error (MAPE) value of $\leq 10\%$. **Results:** The Polar H7 and BSP were valid during both exercise modes (cycling: MAPE = 6.87%, $R = 0.79$; resistance exercise: MAPE = 6.31%, $R = 0.83$). During cycling, the Apple Watch Series 2 revealed the greatest HR validity (MAPE = 4.14%, $R = 0.80$). The BSP revealed the greatest HR accuracy during resistance exercise (MAPE = 6.24%, $R = 0.86$). Across all devices, as exercise intensity increased, there was greater underestimation of HR. No device was valid for EE during cycling or resistance exercise. **Conclusions:** HR from wearable devices differed at different exercise intensities; EE estimates from wearable devices were inaccurate. Wearable devices are not medical devices, and users should be cautious when using these devices for monitoring physiological responses to exercise. **Key Words:** CONSUMER WEARABLES, ACCURACY, HEART RATE, CALORIC EXPENDITURE, AEROBIC TRAINING, WEIGHT TRAINING

Physical inactivity has become a global problem. Recent statistics from the World Health Organization indicate that more than 25% of adults do not meet the recommended guidelines for physical activity and that physical inactivity is the fourth leading cause of human mortality (1). Since the introduction of activity-tracking-wearable-technology devices, monitoring of physical activity levels has become a new phenomenon for consumers and researchers. High usage of wearable technology to monitor physical activity and exercise intensity has been confirmed by a World Wide Survey of Fitness Trends ranking wearable technology first in 2016 and 2017 and is projected to continue as a top

fitness trend (2,3). In a review by Coughlin and Stewart (4), it was revealed that wearable devices can be of benefit to users by increasing their physical activity levels and enhancing weight loss. It was concluded that most existing investigations contained small sample sizes, and it was recommended that future studies assess physical activity using wearable devices in clinical health trials. Wright et al. (5) recently described groundbreaking opportunities for researchers to use consumer activity monitors to conduct physiological research, and some researchers have already begun implementing the use of wearable technology devices in physical activity interventions.

Although tracking physical activity with wearable technology has revealed benefits for users (6,7), a major concern for consumers and researchers is that the continuous feedback from a device is accurate. Two notable features recorded by many wearable devices that are of particular interest to consumers and researchers are caloric (energy) expenditure (EE) and heart rate (HR). EE is of particular relevance to those seeking to accomplish weight management goals (8,9) by following a prescription for exercise volume (10), and accuracy of HR assessment during physical activity is important to properly monitor exercise intensity (10). Over the past decade, technology manufactures have released new models and discontinued existing models; thus, research concerning the accuracy of wearable fitness devices is dynamic.

Address for correspondence: Benjamin Donald Boudreaux, M.S., Department of Kinesiology and Health Studies, Southeastern Louisiana University, 400 Tennessee Ave, Hammond, LA, 70402; E-mail: Bendboudreaux@gmail.com. Submitted for publication June 2017.

Accepted for publication October 2017.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.acsm-msse.org).

0195-9131/18/5003-0624/0

MEDICINE & SCIENCE IN SPORTS & EXERCISE®

Copyright © 2017 by the American College of Sports Medicine

DOI: 10.1249/MSS.0000000000001471

Bai et al. (11) measured the EE of 52 participants during physical activity from seven wearable devices: Fitbit Flex, Jawbone Up 24, Misfit Shine, Nike+ Fuel Band SE, Polar Loop, ActiGraph GT3X+, and BodyMedia Core. Participants performed 25 min of both resistance and aerobic exercise at a self-selected intensity by each subject. EE recorded by wearable devices was compared with that assessed via a metabolic analysis system. The findings suggested that during aerobic exercise, the wearable devices had lower accuracy for EE when compared with a metabolic analysis system. During the unstructured resistance exercise protocol in which participants selected exercises and loads, EE measures from all wearable devices were inaccurate, with mean absolute percent error (MAPE) values greater than or equal to 25%, and were thus considered invalid for determining EE during resistance training. Overall, the wearable devices were inaccurate when measuring EE. It is also important to mention that none of these activity monitors that were investigated measured HR.

Horton et al. (12) recently assessed validity of HR only using the Polar M600 when compared with a three-lead ECG. The protocol consisted of 76 min of different aerobic and resistance exercises at different intensities. Subjects cycled at the four instructed workloads (100 W, 125 W, 150 W, 175 W) and performed treadmill exercise. In the circuit weight training protocol, subjects completed shoulder shrugs, squats, bicep curls, and lunges with dumbbells at a self-selected resistance. Results revealed that the device was mostly accurate during cycling (91.8%) and the least accurate during resistance exercise (34.5%). Moreover, the investigators suggested that future studies should use heterogeneous samples to investigate the effects of 1) different exercise intensities and 2) upper and lower body resistance exercise, on accuracy of wearable devices in which motion artifact and device attachment could affect measurements. This was only the second study to investigate the accuracy of HR during resistance exercise.

Presently, the three studies (11–13) that have investigated the accuracy of wearable devices during resistance exercise have used unstructured, subject-selected resistance exercise intensities. It is important to determine accuracy of wearable devices during resistance exercise regimens structured for individual strength because this form of exercise is necessary to meet the American College of Sports Medicine recommendations for exercise prescription (10). Moreover, no previous study has determined the accuracy of wearable devices during a graded cycling exercise protocol in which revolutions were maintained in a standardized regimen.

The present study was designed to determine the validity of HR and EE of multiple wearable devices during 1) a graded cycling exercise test at constant revolutions per minute with increasing exercise intensities and 2) a standardized regimen including both upper and lower body resistance exercises. In addition, the study was designed to include both male and female subjects of different body compositions. Eight wearable devices (Apple Watch Series 2 (AWS2), Fitbit Blaze (FB), Fitbit Charge 2 (FC2), Garmin Viviosmart HR (GVHR), TomTom Touch (TT), Polar A360 (PA360), Polar H7 (PH7), and

Bose SoundSport Pulse (BSP) headphones) were compared with the respective “gold standard” for measuring HR (ECG) and EE (metabolic analyzer) during exercises that included a graded exercise test on a cycle ergometer (14), and during a resistance exercise trial that included three sets of four resistance exercises performed at a 10-repetition maximum (10-RM) load (15). The first hypothesis was that HR from AWS2 and PH7 would have strong correlations and low MAPE values throughout graded exercise on the cycle ergometer when compared with the ECG. It was expected that other devices including the FB, FC2, GVHR, PA360, TT, and BSP headphones would have weaker validity, and that accuracy of HR assessments from the wearable devices would diminish as exercise intensity increased. The second hypothesis was that all eight wearable devices would generate less accurate HR measurements during resistance exercise than during graded cycling. The third hypothesis was that all wearable devices would reveal inaccurate EE readings during graded cycling and resistance exercise.

METHODS

Participants. Fifty participants (28 women and 22 men) of varying fitness levels between the ages of 18 and 35 yr volunteered for the study. Before participation, eligibility was determined by a brief medical history form to exclude individuals with cardiovascular disease or musculoskeletal injury within the past 6 months. Subjects who were eligible provided written informed consent. The study was approved by the university institutional review board. The mean (\pm SD) age, height, weight, and body mass index of the female subjects were 22.71 ± 2.99 yr, 162.71 ± 5.79 cm, 67.79 ± 14.01 kg, and 25.83 ± 4.83 kg·m⁻². The mean (\pm SD) age, height, weight, and body mass index of the male subjects were 22.00 ± 2.67 yr, 180.14 ± 6.51 cm, 88.55 ± 15.12 kg, and 27.14 ± 3.62 kg·m⁻².

Wearable devices. During the study, subjects wore eight wearable devices (six wrist-worn, one chest-worn, and one ear-worn) simultaneously.

The AWS2 (Apple Inc, Cupertino, CA) is a wrist-worn smartwatch compatible with an iPhone 5® or newer iPhone models with Bluetooth technology for data syncing between the Apple Watch and Activity application. The activity features on this device are step counting, distance tracking, calories, HR, minutes of brisk activity, stand reminders, GPS tracking, and swim laps.

The FB (Fitbit Inc, San Francisco, CA) is a wrist-worn activity tracker compatible with both iOS and Android platforms, which has Bluetooth technology for data syncing with the Fitbit application. Activity features include steps, distance, calories, active minutes, stand reminders, and HR.

The FC2 (Fitbit Inc) is a wrist-worn activity tracker compatible with both iOS and Android platforms, with additional Bluetooth technology for data syncing with the Fitbit application. Activity tracking features include steps, distance, calories,

HR, active minutes, standing reminders, and maximal oxygen uptake estimations.

The PH7 Chest Strap (Polar Electro, Kemple, Finland) is a chest-worn HR monitor that is compatible with both iOS and Android platforms, and requires continuous Bluetooth connection for HR readings from the Polar Beat application. The device's features include HR monitoring and EE.

The PA360 (Polar Electro) is a wrist-worn activity tracker that is compatible with both iOS and Android platforms, and requires Bluetooth connection for data syncing to the Polar Flow application. The device tracks steps, calories, distance, and HR, and provides stand reminders.

The GVHR (Garmin International Inc, Canton of Schaffhausen, Switzerland) is a wrist-worn activity tracker compatible with both iOS and Android platforms, and requires Bluetooth connection for activity data syncing to the Garmin Connect application. Some of the activity tracking features on this device include steps, distance, EE, HR, intensity minutes, and stand reminders.

The TT (TomTom, Amsterdam, the Netherlands) a wrist-worn activity tracker, is compatible with iOS and Android, and requires a Bluetooth connection for activity data syncing with the TomTom Sports application. Some activity tracking features of this device include steps, distance, calories, HR, active minutes, and body composition estimation.

The BSP headphones (Bose Corporation., Framingham, MA) are Bluetooth wireless headphones that were released in September of 2016. This device is compatible with both iOS and Android devices that sync real time HR data via Bluetooth to the Bose Connect application. These headphones are wireless and have only one fitness feature, HR monitoring.

Protocol. All subjects completed two laboratory sessions. At the beginning of each session, height, weight, sex, date of birth, and wrist placement were used to initialize the wearable devices for each subject. After the device set up, a waiting period of 1 to 2 min was allowed for Bluetooth and Wi-Fi or cellular connection with an iPhone 7 Plus (Model A1784; Apple Inc) for demographic synchronization. To prevent bias, placement of wrist-worn devices (three devices on each wrist) was randomized and documented, and placement followed the manufacture's guidelines. The BSP monitored HR only and was not connected for music playback. After subjects were fitted with all eight wearable devices, ECG electrode skin sites were shaved and cleansed with an alcohol wipe. Subjects were then connected to a six-lead ECG (Quinton 4500, Milwaukee, WI) for determining HR and a metabolic analyzer (TrueOne-2400; ParvoMedics, Sandy, UT, USA) with a fitted mask for measuring EE.

During the first session, subjects performed a graded exercise test on a cycle ergometer (Monark, Ergomedic 828E). The protocol began with a 5-min rest period, followed by an HR reading. Next, subjects began the graded exercise test consisting of 2-min stages at 50 rpm, beginning at $300 \text{ kpm}\cdot\text{min}^{-1}$ and increasing by $150 \text{ kpm}\cdot\text{min}^{-1}$ until exhaustion, followed by a 5-min cool down (14). HR was continuously monitored throughout the protocol, but recorded at the end of each initial

phase before increasing flywheel resistance. The HR readings from all wearable devices were digitally time stamped to an iPhone 7 Plus in the Apple Health application and/or in the device's specific application including the BSP. HR was recorded from the ECG at each time point and confirmed by measuring the distance between R and R waves in consecutive cadence cycles from hardcopy ECG printouts. EE from seven wearable devices, excluding the BSP, was determined at the end of the protocol and compared with that from the metabolic analyzer.

One hour after the graded exercise test, a 10-RM was determined for four different strength training exercises performed on a resistance exercise machine (BK 620 Super Jungle; Body Masters, Rayne, LA): two upper body exercises (chest press, latissimus dorsi (lat) pulldown) and two lower body exercises (leg extension and leg curl) (15). Subjects were instructed to warm up with 5–10 repetitions at a 40%–60% perceived maximal exertion. After the warm-up, they performed a set of 10 repetitions using a resistance perceived to be their 10-RM load. If the 10th repetition was not achieved, or more than 10 repetitions were completed, a second trial was performed using more or less weight. Five minutes of rest between attempts occurred, and the 10-RM attempts were alternated between upper body and lower body exercises. Finally, subjects performed a second test for verification that the 10-RM resistance load was reliable (15).

Subjects reported for the second session within a 3-d time frame. As in session 1, wearable devices were placed on each subject and initialized using their demographic characteristics, and subjects were connected to a six-lead ECG and metabolic analyzer. Subjects then performed three circuits of 10 repetitions at the previously determined 10-RM for each of the four exercises in the following order: leg curl, chest press, leg extension, and lat pulldown. After completing each exercise, subjects remained seated for 7 s to receive clear HR readings from the ECG and wearable devices before moving to the next exercise. All exercise circuits and repetitions were performed to a 2-s lifting and 2-s lowering cadence (15) emitted from the iPhone 7 Plus and played through the BSP at a volume between 55 and 65 dB. This process standardized the exercise speed among subjects. The same protocol for determining HR and EE was used across all devices as in the first session.

Statistical analyses. HR and EE data for wearable devices were retrieved from the iPhone 7 Plus during the cycling and resistance exercise. Data were analyzed using Version 20.0 of SPSS (IBM Corp, Somers, NY). The resultant data included HR at 6 time points during graded cycling, and 14 time points during resistance training. It also included the average HR during each session and EE measurements at the completion of the cycling and strength training protocols. Four statistical procedures were used to examine validity of wearable device measurements. Data from wearable devices were compared with the ECG and metabolic analyzer using paired *t*-tests, with the *P* value of 0.05 used as the threshold for significant differences. Intraclass correlation was calculated to examine relationships between values from each wearable

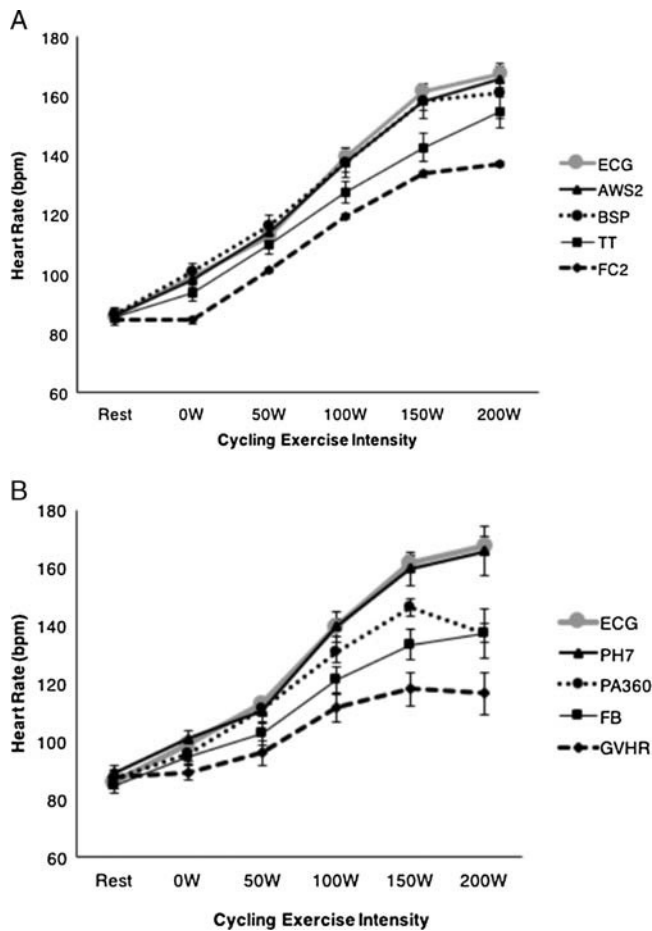


FIGURE 1—Values represent HR (mean ± SE) during graded exercise cycling for ECG and wearable devices. A, ECG, AWS2, BSP, TT, and FC2. B, ECG, PH7, PA360, FB, and GVHR.

device and its corresponding gold standard. For the purpose of validity classification, the intraclass correlation thresholds suggested by Fokkema et al. (16) were used: excellent, ≥ 0.90 ; good, 0.75–0.90; moderate, 0.60–0.75; and low, ≤ 0.60 . Third, MAPE, representing the error percentage between measures was calculated. MAPE does not have a standardized threshold for determination of accuracy/validity of measurements.

However, Fokkema et al. (16) suggest a MAPE threshold of $\leq 5\%$, whereas Nelson et al. (17) used a MAPE threshold of $\leq 10\%$ to classify a wearable device as valid. The $\leq 10\%$ MAPE value was used in the present study as the criterion measure for validity. The final statistical assessment conducted was a Bland–Altman analysis that included the mean difference and upper/lower limits of HR from the wearable devices compared with the ECG during cycling and resistance exercise (see Tables, Supplemental Digital Content 1, Mean differences and 95% confidence intervals revealing gradual increases in the mean differences and 95% confidence intervals from continuous HR and underestimations for five devices and diverse mean differences and 95% confidence intervals for three devices, <http://links.lww.com/MSS/B89>; Supplemental Digital Content 2, Mean differences and 95% confidence intervals revealing wrist-worn devices underestimated HR throughout each resistance exercise workload, and two devices maintained more accurate HR readings that both over and underestimated ECG values, <http://links.lww.com/MSS/B90>).

RESULTS

HR during cycling. Figure 1 displays the mean and SE values for HR at each time point during graded cycling for ECG and wearable devices. Table 1 provides corresponding intraclass correlation coefficients and MAPE values for each wearable device with indications of when HR from wearable devices were significantly different from ECG readings.

Intraclass correlation values for HR from the various wearable devices and the ECG were diverse and were stronger at rest and decreased as exercise intensity increased. At rest, HR from the most wearable devices had strong relationships to ECG values ($R = 0.76$ – 0.99). When exercise began as well as during each increase in exercise intensity, intraclass correlation coefficients were reduced in most devices (e.g., $R = 0.47$ – 0.90 at 0 W; $R = 0.32$ – 0.85 at 100 W; $R = 0.11$ – 0.80 at 150 W). Among the devices, three (AWS2, PH7, and BSP) maintained “good” correlational values ($R \geq 0.75$) throughout the majority of the cycling protocol and on average during the

TABLE 1. HR during graded exercise cycling for eight wearable devices compared with ECG.

Workload	FB	FC2	TT	AWS2	BSP	PA360	GVHR	PH7
Rest	ICC = 0.92 MAPE = 4.43	ICC = 0.92 MAPE = 4.68	ICC = 0.97 MAPE = 3.53	ICC = 0.99 MAPE = 1.21	ICC = 0.97 MAPE = 3.24	ICC = 0.85 MAPE = 7.56	ICC = 0.76 MAPE = 7.30	ICC = 0.91 MAPE = 5.47*
0 W	ICC = 0.70 MAPE = 10.08*	ICC = 0.69 MAPE = 10.55*	ICC = 0.53 MAPE = 13.82*	ICC = 0.87 MAPE = 4.40	ICC = 0.90 MAPE = 4.63	ICC = 0.66 MAPE = 14.7	ICC = 0.47 MAPE = 16.70*	ICC = 0.90 MAPE = 6.71
50 W	ICC = 0.47 MAPE = 19.75*	ICC = 0.47 MAPE = 20.75*	ICC = 0.74 MAPE = 8.54	ICC = 0.90 MAPE = 2.99	ICC = 0.75 MAPE = 6.40	ICC = 0.52 MAPE = 19.59	ICC = 0.34 MAPE = 28.67*	ICC = 0.79 MAPE = 5.94*
100 W	ICC = 0.48 MAPE = 22.90*	ICC = 0.47 MAPE = 21.74*	ICC = 0.56 MAPE = 12.86*	ICC = 0.85 MAPE = 4.84	ICC = 0.78 MAPE = 8.26	ICC = 0.55 MAPE = 17.88	ICC = 0.32 MAPE = 27.75*	ICC = 0.77 MAPE = 7.78
150 W	ICC = 0.31 MAPE = 30.97*	ICC = 0.32 MAPE = 31.38*	ICC = 0.30 MAPE = 20.93*	ICC = 0.80 MAPE = 4.26	ICC = 0.76 MAPE = 6.70	ICC = 0.42 MAPE = 21.98*	ICC = 0.11 MAPE = 38.73*	ICC = 0.63 MAPE = 6.94
200 W	ICC = 0.12 MAPE = 38.24*	ICC = 0.14 MAPE = 39.05*	ICC = 0.46 MAPE = 14.31*	ICC = 0.97 MAPE = 7.16*	ICC = 0.50 MAPE = 15.42	ICC = 0.11 MAPE = 35.11*	ICC = 0.01 MAPE = 33.14*	ICC = 0.85 MAPE = 8.39
Average	ICC = 0.50 MAPE = 21.06	ICC = 0.58 MAPE = 21.36	ICC = 0.59 MAPE = 12.33	ICC = 0.90 MAPE = 4.14	ICC = 0.78 MAPE = 7.44	ICC = 0.52 MAPE = 19.48	ICC = 0.36 MAPE = 25.38	ICC = 0.79 MAPE = 6.87

Validity was established as having a MAPE value of $\leq 10\%$.
Sample size contained 25 participants for the 200-W stage.
*Significantly different from ECG ($P < 0.05$).
ICC, intraclass correlation coefficient.

session. HR from the remaining wearable devices tended to have low ($R \leq 0.60$) intraclass correlation with ECG.

MAPE values followed a similar trend, reflecting lower levels of error during rest (1.21%–7.56%) and higher levels of error as exercise intensity increased (e.g., 4.40%–16.70% at 0 W; 4.84%–27.75% at 100 W). Two wearable devices, AWS2 and PH7, maintained MAPE values of <10% through all levels, and the BSP maintained this MAPE criterion during five of six time points. The remaining wearable devices had MAPE values of 10% or higher at all or most stages.

As depicted in Figure 1, HR measurements from ECG and wearable devices were similar at rest, but as exercise intensity increased, disparity from ECG increased in many devices. The results of *t*-tests comparing HR from devices with ECG at each workload indicated that the BSP was the only device that provided HR readings that were not significantly different from the ECG readings at all workloads of the cycling protocol. The AWS2 was not significantly different from the ECG for all but one workload (200 W), and the PH7 for all but two time points. Excluding rest, three wearable devices had HR values that were significantly different from the ECG at all stages (FB, FC2, and GVHR).

Bland–Altman analysis revealed gradual increases in the mean differences and 95% confidence intervals from continuous HR underestimations for five of the eight devices (FB, FC2, GVHR, PA360, TT) as exercise intensity progressed. The three remaining devices (AWS2, BSP, PH7) were diverse in the mean differences and 95% confidence intervals, with slightly overestimated HR values at rest or low cycling intensities and slightly underestimated HR values at higher cycling intensities. Differences in HR readings from ECG were lowest for BSP, AWS2, and PH7 (see Table, Supplemental Digital Content 1, Mean differences and 95% confidence intervals revealing gradual increases in the mean differences and 95% confidence intervals from continuous HR and underestimations for five devices and diverse mean differences and 95% confidence intervals for three devices, <http://links.lww.com/MSS/B89>).

Caloric expenditure during cycling. Table 2 provides the mean and SE for EE from wearable devices and the metabolic analyzer during the cycling session, as well as intraclass correlation coefficients and MAPE values. When compared using dependent *t*-tests, only the GVHR caloric average was not significantly different from the metabolic analyzer. Caloric values from all seven devices had weak

correlational relationships to the metabolic analyzer, and also displayed high MAPE values. The FC2 had the weakest correlation ($R = 0.18$) and the highest MAPE of any wearable (75.15%). The wearable that had the strongest correlation with the metabolic analyzer was the GVHR ($R = 0.41$), and the AWS2 had the lowest MAPE at 21.13%. There was a tendency for some devices to consistently overpredict or underpredict EE. To reflect this, the number of subjects whose EE was overpredicted and underperdicted by each device is reported (Table 2). The AWS2 overestimated EE in 49 of the 50 subjects, whereas the FB underestimated EE (41 of 50 subjects) more often. By comparison, the numbers of overestimated and underestimated values (18,19) for the GVHR were relatively similar.

HR during resistance exercise. Intraclass correlations, MAPE values with indications of significant differences for HR between wearable devices, and ECG during resistance exercise are presented in Table 3. During resistance exercise, the wearable devices had diverse interclass correlation values that tended to decline from rest to increasing resistance exercise volume (e.g., $R = 0.50$ – 0.99 at rest; $R = 0.16$ – 0.82 at chest press 1—exercise 2, $R = 0.08$ – 0.82 at lat pulldown 3—exercise 12). Similarly, increases in MAPE values were observed as resistance exercise volume increased over the training regimen, ranging from 1.44% to 9.97% at rest, and 5.47% to 21.20% at the completion of the last exercise. The two non-wrist-worn devices, PH7 and BSP, demonstrated strongest HR validity during strength training, with an average correlation to ECG at 0.80 or higher and MAPE values of $\leq 10\%$ (PH7: $R = 0.83$, MAPE = 6.31%; BSP: $R = 0.86$, MAPE = 6.24%). Validity was lower for wrist-worn devices, with PA360 ($R = 0.68$, MAPE = 8.66%), AWS2 ($R = 0.72$, MAPE = 10.99%), and FC2 ($R = 0.59$, MAPE = 9.97%) demonstrating moderate correlations and lower error values than other devices.

Figure 2 depicts HR measured by the eight wearable devices and ECG throughout the entire resistance exercise protocol. Changes in HR during strength training from BSP, PH7, PA360, and AWS2 tended to mirror ECG values, whereas values from FC2, TT, FB, and GVHR varied from and were lower than those from the ECG. During strength training, HR readings from all eight wearable devices were significantly different from the ECG at least once during the 12 HR measurements (see Table 3). From most accurate to least

TABLE 2. Caloric expenditure during graded exercise cycling for seven wearable devices compared with metabolic analysis.

Device	Mean \pm SE	ICC (<i>R</i>)	MAPE, %	No. of Underestimations	No. of Overestimations
Metabolic analyzer	64.48 \pm 3.64				
FC2	37.28 \pm 3.56*	0.18	75.15	40	10
FB	39.14 \pm 4.10*	0.20	72.01	41	8
GVHR	69.90 \pm 5.81	0.41	63.05	24	26
PA360	83.36 \pm 4.94*	0.28	38.18	13	37
TT	95.92 \pm 6.93*	0.30	41.27	10	40
PH7	97.62 \pm 5.35*	0.27	29.83	7	43
AWS2	128.72 \pm 6.34*	0.23	21.13	1	49

Validity was established as having a MAPE value of $\leq 10\%$.

*Significantly different from metabolic analysis ($P < 0.05$).

ICC, intraclass correlation coefficient.

TABLE 3. HR during resistance exercise in eight wearable devices compared to ECG.

Exercise	FC2	FB	TT	AWS2	GVHR	PA360	BSP	PH7
Rest	ICC = 0.83 MAPE = 6.14	ICC = 0.93 MAPE = 4.00	ICC = 0.89 MAPE = 5.95*	ICC = 0.99 MAPE = 1.44	ICC = 0.82 MAPE = 7.26*	ICC = 0.67 MAPE = 10.00*	ICC = 0.79 MAPE = 8.52*	ICC = 0.50 MAPE = 9.97*
LC 1	ICC = 0.72 MAPE = 8.35	ICC = 0.43 MAPE = 33.20*	ICC = 0.42 MAPE = 13.18*	ICC = 0.75 MAPE = 6.46	ICC = 0.45 MAPE = 10.36*	ICC = 0.68 MAPE = 10.36	ICC = 0.83 MAPE = 5.97	ICC = 0.75 MAPE = 8.35
CP 1	ICC = 0.80 MAPE = 8.08	ICC = 0.53 MAPE = 31.96	ICC = 0.16 MAPE = 19.00*	ICC = 0.78 MAPE = 9.76	ICC = 0.49 MAPE = 9.40	ICC = 0.66 MAPE = 7.39*	ICC = 0.82 MAPE = 8.44*	ICC = 0.74 MAPE = 8.10*
LE 1	ICC = 0.75 MAPE = 7.46*	ICC = 0.77 MAPE = 8.08*	ICC = 0.07 MAPE = 19.37*	ICC = 0.70 MAPE = 10.78	ICC = 0.39 MAPE = 9.42*	ICC = 0.70 MAPE = 8.71	ICC = 0.85 MAPE = 6.67	ICC = 0.83 MAPE = 6.42*
LPD 1	ICC = 0.69 MAPE = 9.20*	ICC = 0.67 MAPE = 10.43*	ICC = 0.02 MAPE = 19.35*	ICC = 0.76 MAPE = 8.60	ICC = 0.47 MAPE = 8.40*	ICC = 0.71 MAPE = 8.64*	ICC = 0.88 MAPE = 6.25	ICC = 0.92 MAPE = 5.63
LC 2	ICC = 0.60 MAPE = 8.12*	ICC = 0.59 MAPE = 10.15*	ICC = 0.06 MAPE = 20.96*	ICC = 0.55 MAPE = 15.62*	ICC = 0.53 MAPE = 8.19	ICC = 0.48 MAPE = 9.40*	ICC = 0.85 MAPE = 6.48	ICC = 0.90 MAPE = 6.17
CP 2	ICC = 0.69 MAPE = 8.94*	ICC = 0.69 MAPE = 9.76*	ICC = 0.02 MAPE = 19.65*	ICC = 0.65 MAPE = 16.21*	ICC = 0.68 MAPE = 8.48*	ICC = 0.68 MAPE = 7.94	ICC = 0.87 MAPE = 6.26	ICC = 0.87 MAPE = 4.63*
LE	ICC = 0.54 MAPE = 9.74*	ICC = 0.55 MAPE = 10.53*	ICC = 0.08 MAPE = 22.46*	ICC = 0.63 MAPE = 12.67*	ICC = 0.37 MAPE = 8.60*	ICC = 0.63 MAPE = 9.56*	ICC = 0.95 MAPE = 3.74*	ICC = 0.92 MAPE = 4.81*
LPD 2	ICC = 0.44 MAPE = 12.06*	ICC = 0.47 MAPE = 11.73*	ICC = 0.06 MAPE = 22.59*	ICC = 0.67 MAPE = 8.47*	ICC = 0.46 MAPE = 9.15*	ICC = 0.67 MAPE = 8.14*	ICC = 0.87 MAPE = 4.92*	ICC = 0.91 MAPE = 4.64
LC 3	ICC = 0.29 MAPE = 14.73*	ICC = 0.34 MAPE = 13.89*	ICC = 0.09 MAPE = 19.18*	ICC = 0.39 MAPE = 21.20*	ICC = 0.42 MAPE = 10.84*	ICC = 0.71 MAPE = 5.32*	ICC = 0.90 MAPE = 4.46	ICC = 0.90 MAPE = 8.08
CP 3	ICC = 0.42 MAPE = 13.86*	ICC = 0.56 MAPE = 10.67*	ICC = 0.05 MAPE = 22.63*	ICC = 0.62 MAPE = 13.69*	ICC = 0.51 MAPE = 10.68*	ICC = 0.75 MAPE = 7.25	ICC = 0.91 MAPE = 5.28*	ICC = 0.88 MAPE = 5.23
LE 3	ICC = 0.49 MAPE = 12.01*	ICC = 0.47 MAPE = 12.60*	ICC = 0.02 MAPE = 23.25*	ICC = 0.56 MAPE = 12.56*	ICC = 0.26 MAPE = 25.46*	ICC = 0.75 MAPE = 6.36*	ICC = 0.92 MAPE = 4.05*	ICC = 0.87 MAPE = 4.03
LPD 3	ICC = 0.38 MAPE = 10.96*	ICC = 0.37 MAPE = 11.67*	ICC = 0.08 MAPE = 21.20*	ICC = 0.77 MAPE = 5.47*	ICC = 0.17 MAPE = 12.30*	ICC = 0.61 MAPE = 8.91*	ICC = 0.72 MAPE = 10.10*	ICC = 0.82 MAPE = 6.56*
Average	ICC = 0.59 MAPE = 9.97	ICC = 0.57 MAPE = 13.74	ICC = 0.16 MAPE = 19.14	ICC = 0.72 MAPE = 10.99	ICC = 0.46 MAPE = 10.66	ICC = 0.68 MAPE = 8.66	ICC = 0.86 MAPE = 6.24	ICC = 0.83 MAPE = 6.31

Validity was established as having a MAPE value of $\leq 10\%$.

*Significantly different from ECG ($P < 0.05$).

CP, chest press; ICC, intraclass correlation coefficient; LC, leg curl; LE, leg extension; LPD, lat pull down.

accurate, the number of time points devices' readings were not significantly different from ECG during exercise were as follows: PH7, 8 of 12 measurements; BSP, 7 of 12; PA360, 4 of 12; AWS2, 4 of 12; GVHR, 2 of 12; FC2, 2 of 12; FB, 1 of 12; and TT, 0 of 12.

Bland-Altman analysis revealed diverse values in the mean difference and 95% confidence intervals across all eight wearable devices compared with the ECG. All six wrist-worn devices (AWS2, FB, FC2, GVHR, PA360, TT) underestimated HR throughout every resistance exercise workload except at rest. The two remaining devices (BSP, PH7) maintained relatively closer HR values to ECG after resistance exercises, with both overestimation and underestimation (see Table, Supplemental Digital Content 2, Mean differences and 95% confidence intervals revealing wrist-worn devices underestimated HR throughout each resistance exercise workload, and two devices maintained more accurate HR readings that both over and underestimated ECG values, <http://links.lww.com/MSS/B90>).

Caloric expenditure during resistance exercise.

Descriptive statistics for EE and the results of analyses are presented in Table 4. EE from all wearable devices had weak intraclass correlational relationships (0.02–0.18) compared with the metabolic analyzer, with the strongest correlation ($R = 0.18$) shared by the AWS2 and GVHR. The TT had the lowest correlational value ($R = 0.02$). In addition, caloric estimations from all wearable devices had high MAPE values (42.69%–57.02%) compared with the metabolic analyzer. The AWS2 had the lowest MAPE (42.69%), whereas the GVHR had the highest (57.02%). When each subject's data were examined, all wearable devices tended to overestimate EE during resistance exercise. For example, the AWS2 overestimated EE

for 45 of 50 subjects, PH7 overestimated for 43 of 50, and FC2 overestimated for 30 of 50 subjects.

DISCUSSION

The present study examined the validity of HR and EE estimations of various wearable devices during a cycle ergometer graded exercise test and during a structured resistance exercise session. The findings of this study revealed that both HR and EE differed among the eight wearable devices during both cycling and resistance exercise, and had varying levels of validity when compared with a six-lead ECG and metabolic analyzer. It was also observed that HR measures from wearable devices were more accurate at rest and lower exercise intensities than at higher intensities. Among tested devices, HR accuracy, as reflected by intraclass correlation and MAPE values, was highest in the PH7, BSP, and AWS2. The PH7 and AWS2 also proved to provide more accurate caloric estimations than other devices. This is the first study to determine the accuracy of wearable devices during a graded cycling exercise test and during a structured resistance exercise regimen.

The first hypothesis, which states that two devices, AWS2 and PH7, would have strong correlational values and low MAPE values throughout the entire graded exercise protocol on the cycle ergometer compared with the 6 lead ECG, was supported. Relative to other wearable devices, the validity of HR readings was strongest for the AWS2 ($R = 0.80$ – 0.99 ; $MAPE = 2.99\%$ – 7.16%) and PH7 ($R = 0.67$ – 0.90 ; $MAPE = 5.94\%$ – 8.39%). These results support the findings by Wang et al. (20), who found that the Apple Watch and PH7 had high levels of

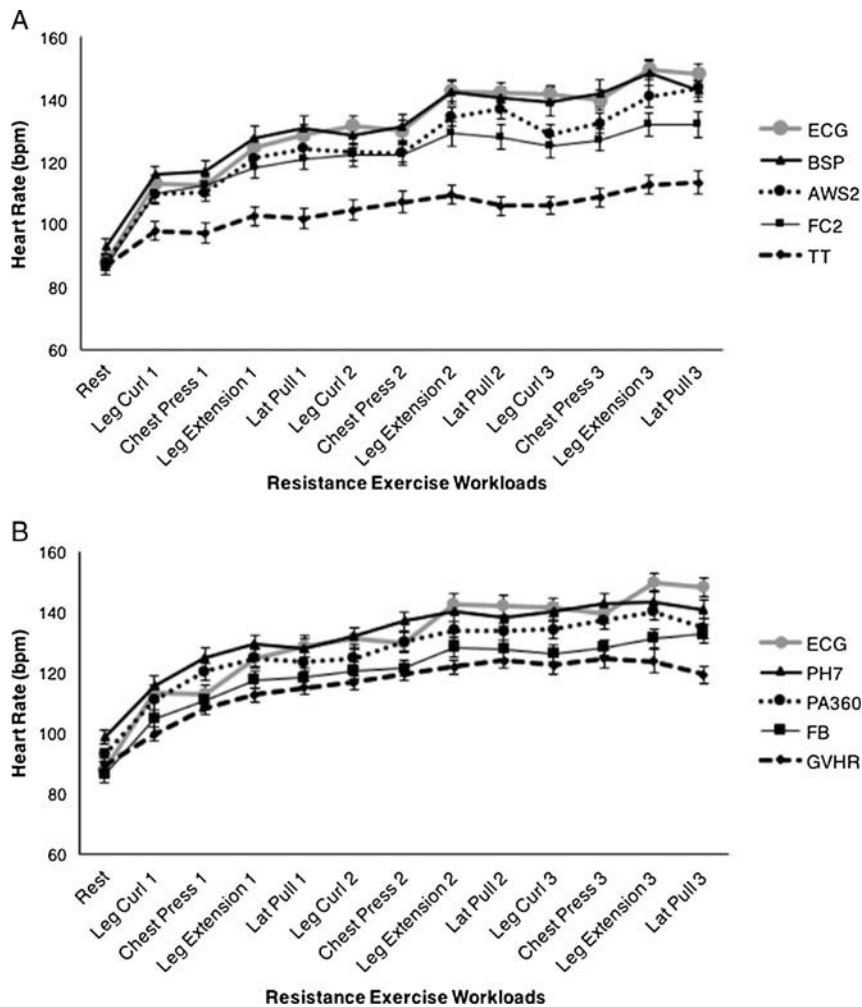


FIGURE 2—Values represent HR (mean ± SE) during resistance exercise for ECG and wearable devices. A, ECG, BSP, AWS2, FC2, and TT. B, PH7, PA360, FB, and GVHR.

accuracy in HR during aerobic activity. The BSP was also found to be a promising device in this protocol, maintaining valid HR readings for most of the cycling protocol ($R = 0.50$ – 0.90 ; $MAPE = 4.63\%$ – 15.42%). The five remaining wearable devices (FB, FC2, PA360, GVHR, TT) were valid for HR at rest and only few stages of cycling. To date, no existing study has investigated the validity of HR measurements in wearable devices from a gradual low to vigorous intensity on a cycle ergometer. Previous researchers (12,13,21–23) investigated HR from wearable devices on a cycle ergometer, but at

selected exercise intensities and sometimes with additional physical activities in one protocol. Shcherbina et al. (22) recently compared HR from seven wearable devices with a 12-lead ECG during two intensities of aerobic activities. Activities included sitting, walking (3.0 mph, 4.0 mph), running (6.0 mph, 7.5 mph), and cycling (100 W, 175 W). Results from the study indicated that the wearable devices had the highest inaccuracies during walking and the lowest inaccuracies during cycling, but varied among the devices during running. The present study revealed that accuracy of wearable

TABLE 4. Caloric expenditure during resistance exercise for seven wearable devices compared with metabolic analysis.

Device	Mean ± SE	ICC (R)	MAPE, %	No. of Underestimations	No. of Overestimations
Metabolic analyzer	74.98 ± 3.41				
FC2	69.20 ± 4.64	0.08	47.85	18	30
FB	71.54 ± 4.52	0.14	49.07	20	30
GVHR	97.82 ± 5.89*	0.18	57.02	12	38
TT	99.58 ± 7.54*	0.02	51.64	18	32
PA360	112.18 ± 6.31*	0.13	52.95	9	41
PH7	132.30 ± 8.96*	0.11	43.29	7	43
AWS2	140.50 ± 7.88*	0.18	42.69	5	45

Validity was established as having a MAPE value of $\leq 10\%$.

*Significantly different from metabolic analysis ($P < 0.05$).

ICC, intraclass correlation coefficient.

devices during a graded exercise test was reduced as workload increased.

The second hypothesis, which states that all eight wearable technology devices would have weak correlations and high MAPE values compared with a six-lead ECG when subjects performed resistance exercises, was not supported. Four of the eight wearable devices (BSP, PH7, PA360, and FC2) had average MAPE values of $\leq 10\%$ and two devices had correlation values of ≥ 0.80 relative to ECG, indicating that they were mostly accurate. The two non-wrist-worn wearables met the criteria for validity ($r \geq 0.75$ and MAPE $\leq 10\%$): BSP (average $R = 0.86$; MAPE = 6.24%) and PH7 (average $R = 0.83$; MAPE = 6.31%). Among wrist-worn devices, the PA360 (average $R = 0.68$; MAPE = 8.66%) and AWS2 (average $R = 0.72$; MAPE = 10.99%) were most accurate; other devices were less valid. In a previous study, Jo et al. (13) determined HR accuracy of the Basis Peak and Fitbit Charge HR during a protocol in which subjects performed dumbbell arm raises and dumbbell lunges within a multiple physical activity protocol. The results of the study revealed that the Basis Peak was accurate during the two resistance exercises.

A comparison of the MAPE values for the different workloads in both graded cycling and resistance exercise indicated that the PH7 and BSP were the most accurate wearable devices overall for both modes of exercise. A comparison of the MAPE values for the different workloads in both graded cycling and resistance exercise indicated that the PH7 and BSP were the most accurate wearable devices overall for both modes of exercise. Among tested devices, the PH7 was the sole device to meet the validity criteria for HR during both graded cycling and resistance exercise for every workload; however, BSP had greater validity than PH7 at some workloads. Other devices found to be more accurate in one exercise mode relative to others were the AWS2 and PA360. The present study is also unique because of its investigation of the BSP consumer Bluetooth wireless headphones that measured HR during both cycling and resistance exercise, and the results support further examination. LeBoeuf et al. (24) conducted the only known study that investigated HR-reading headphones during a cardiopulmonary exercise test. The device found to be most accurate for HR, the PH7, is chest worn, and as stated by Polar, the technology measures the electrical signals generated by the heart for each beat (similar to ECG), whereas wrist-worn devices and ear phones use photoplethysmography (PPG), or a pulse oximeter to measure HR. PPG is a low-cost medical technique applied to the skin that uses the transmission and reflection of light into the skin to measure changes in blood volume within a specific tissue (25). Previous research (18,19,26–28) suggests that PPG may have limitations in measuring HR that arise from the continuous increase and decrease of compression of a wearable device's HR sensor on the skin (18,19,26–28). Tight compression of a wearable device's sensor on the skin can create noise artifacts in the waveform distributions of the device, weakening PPG signals.

Compression in a wearable device creates difficulties for the PPG signals to shine into the skin and receive clear signal changes of blood volume. A second problem arises from changes in skin perfusion as skin temperatures increase at higher exercise intensities. Maeda et al. (25) reported that skin temperatures lower than 20°C or higher than 38°C led to weak PPG signals for measuring HR. A third problem in these wrist-worn devices involves movement artifact with the repeated contraction and relaxation of the skeletal muscles in the forearm and hand. Rafolt and Gallasch (27) reported that movement artifact in a wearable device can be impacted by the repeated pull of gravity on a device, and increased and decreased tension of a wearable device on the skin by the repeated movement of a limb, which may cause artifacts. It is possible that the lack of movement artifact produced against the PPG sensor in the ears when wearing the BSP headphones contributed to the accuracy of the PPG for HR.

The third hypothesis was that all eight wearable technology devices would have inaccurate EE readings, reflected in high MAPE values and low correlation scores compared with a metabolic analyzer during both resistance exercise and graded cycling. This prediction was supported. EE estimates during cycling were diverse among the seven wearables, and analyses yielded high MAPE values (range, 21.13%–75.15%) and low correlational relationships (0.18–0.41) to the metabolic analyzer. In addition, most wearable devices tended to consistently overestimate calories (AWS2, PH7, TT, PA360) or consistently underestimate calories (FB, FC2). Similarly, during the resistance exercise protocol, high MAPE values (42.69%–57.02%) and weak correlational relationships (0.02–0.18) were also observed when measurements from wearable devices were compared with metabolic analysis results.

Findings from this study support previous research (11,22,23,29) that wearable devices may not be accurate for measurement of EE during physical activity. For example, Nelson et al. (17) investigated five wearable devices, both wrist-worn and hip-worn, during a variety of physical activities. Their findings indicated that during ambulatory activities, all wearable devices overestimated EE except for one, which underestimated when compared with a metabolic system. The probable reason for wearable devices overpredicting or underpredicting EE is the application of inaccurate metabolic equations and the use of inaccurate HR to measure exercise intensity by a user. During the initial setup process of a wearable device, the user is prompted to provide certain preliminary demographics, such as sex, date of birth, height, and weight that are likely used in the metabolic equations used by the device to calculate EE. These equations, however, are unknown to both users and researchers. Before the present investigation, no study had investigated a consumer wearable device with an HR sensor and its influence in EE estimation during a structured resistance exercise regimen. It has been shown that the set configuration plays a strong role in affecting the cardiovascular response and fatigue during resistance exercise (30). In addition, it has been shown that there is a relationship between hemodynamic responses and central

fatigue (30). Moreover, there is some evidence that an HR-based recovery period for strength training can be more efficient and effective for hypertrophic strength training (31). Thus, using a wearable device that accurately determines HR could be helpful for training purposes.

The present study provides new insights into the accuracy of wearable devices during physical activity, but device latency or ensemble averaging of HR from wearable devices is a limitation of a study of this type. Some wearable devices did not immediately record HR at the exact time point desired. In wearable devices for which HR was delayed, a reading was taken no more than 3 to 5 s before or after the ECG reading. The present study was unique in that no existing study has investigated the validity of HR from wearable devices in a controlled, standardized resistance exercise session. Another strength is that this is the first study to investigate multiple wearable devices simultaneously during different modes of exercise and different exercise intensities. In addition, all wrist-worn devices were alternated for position on every subject during each exercise mode. Moreover, this is the first study to investigate the accuracy for HR in consumer headphones. Future studies should continue to investigate the validity of wearable devices in HR and EE during physical activity, examine the accuracy of other activity tracking features (e.g., steps, distance), and determine the accuracy of other modalities of physical activity. Examples include a free weight session, high-intensity interval

training on a cycle ergometer, and swimming because some new wearable devices are waterproof.

Wearable technology devices such as smartwatches or activity trackers are a popular fitness trend to promote physical activity, but users should be aware that they are not medical devices, nor are they regulated by the Food and Drug Administration, and the accuracy of measurements during some activities may be low. This study examined the validity of HR and EE in eight wearable devices. Although some devices in this study were valid for determining HR, the readings varied during different forms and intensities of physical activity. Moreover, it was found that the higher the exercise intensity during cycling and resistance exercise, the greater the tendency was for most devices to underpredict HR. Given the findings from the current study, it is clear that EE measures from devices in the present study should be used for estimation purposes only, and this feature should be used with caution. For researchers, the findings point out the more accurate devices for measuring HR during graded cycling and during resistance exercise that could be used for general physiological assessments during an intervention.

No funding support was provided for this study. There are no relevant conflicts of interest to disclose.

The results of the study do not constitute endorsement by the American College of Sports Medicine. The results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

REFERENCES

1. World Health Organization. *Physical Activity Fact Sheet*. [cited 1 Aug 2017]. Available from: <http://www.who.int/mediacentre/factsheets/fs385/en/>.
2. Thompson W. Worldwide survey of fitness trends for 2016: 10th anniversary edition. *ACSM Health Fitness J*. 2015;19(6):9–18.
3. Thompson W. Worldwide survey of fitness trends for 2017. *ACSM Health Fitness J*. 2016;20(6):8–17.
4. Coughlin SS, Stewart J. Use of consumer wearable devices to promote physical activity: a review of health intervention studies. *J Environ Health Sci*. 2016;2(6):1–10.
5. Wright SP, Hall Brown TS, Collier SR, Sandberg K. How consumer activity monitors could transform human physiology research. *Am J Physiol Regul Integr Comp Physiol*. 2017;312(3):R358–267.
6. Ellingson LD, Meyer JD, Cook DB. Wearable technology reduces prolonged bouts of sedentary behavior. *Translat J ACSM*. 2016;1(2):10–7.
7. Webber SC, Strachan SM, Pachu NS. Sedentary behavior, cadence and physical activity outcomes after knee arthroplasty. *Med Sci Sports Exerc*. 2017;49(6):1057–65.
8. Donnelly JE, Blair SN, Jakicic JM, Manore MM, Rankin JW, Smith BK. Appropriate physical activity intervention strategies for weight loss and prevention of weight regain for adults. *Med Sci Sports Exerc*. 2009;41(2):459–71.
9. Manore MM, Brown K, Houtkooper L, et al. Energy balance at crossroads: translating the science into action. *Med Sci Sports Exerc*. 2014;46(7):1466–73.
10. American College of Sports Medicine. *ACSM's Guidelines for Exercise Testing and Prescription*. 10th ed. Philadelphia (PA): Wolters Kluwer; 2018.
11. Bai Y, Welk GJ, Nam YH, et al. Comparison of consumer and research monitors under semistructured settings. *Med Sci Sports Exerc*. 2015;48(1):151–8.
12. Horton JF, Stergiou P, Fung TS, Katz L. Comparison of Polar M600 optical heart rate and ECG heart rate during exercise. *Med Sci Sports Exerc*. 2017;49(12):2600–7.
13. Jo E, Lewis K, Directo D, Kim MJ, Dolezal BA. Validation of biofeedback wearable devices for photoplethysmographic heart rate tracking. *J Sports Sci Med*. 2016;15(3):540–7.
14. Hetrick MM, Naquin M, Gillan WW, Williams BM, Kraemer RR. A hydrothermally processed maize starch and its effects on blood glucose levels during high intensity interval exercise. *J Strength Cond Res*. 2018;32(1):3–12.
15. Durand RJ, Castracane VD, Hollander DB, et al. Hormonal responses from concentric and eccentric muscle contractions. *Med Sci Sports Exerc*. 2003;35(6):937–43.
16. Fokkema T, Kooiman TJ, Krijnen WP, Cees P, Schans VD, Groot MD. Reliability and validity of ten consumer activity trackers depend on walking speed. *Med Sci Sports Exerc*. 2017;49(4):793–800.
17. 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):1619–29.
18. Achten J, JueKendrup AE. Heart rate monitoring applications and limitations. *Sports Med*. 2003;33(7):517–38.
19. Jeong C, Yoon H, Kang H, Yeom H. Effects of skin surface temperature on photoplethysmograph. *J Healthc Eng*. 2014;5(4):429–38.
20. Wang R, Blackburn G, Desai M, et al. Accuracy of wrist-worn heart rate monitors. *JAMA Cardiol*. 2016;2(1):104–6.

21. 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.
22. 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). pii: E3. doi: 10.3390/jpm7020003.
23. Wallen MP, Gomersall SR, Keating SE, Wisløff U, Coombes JS. Accuracy of heart rate watches: implications for weight management. *PLoS One.* 2016;1–11.
24. LeBoeuf SF, Aumer ME, Kraus WE, Johnson JL, Duscha B. Earbud-based sensor for the assessment of energy expenditure, heart rate, and VO₂max. *Med Sci Sports Exerc.* 2014;46(5):1046–52.
25. Maeda Y, Sekine M, Tamura T. The advantages of wearable green reflected photoplethysmography. *J Med Syst.* 2011;35(5):829–34.
26. Butler MJ, Crowe JA, Hayes-Gill BR, Rodmell PI. Motion limitations of non-contact photoplethysmography due to the optical and topological properties of skin. *Physiol Meas.* 2016;37(5):N27–37.
27. Rafolt D, Gallasch E. Influence of contact forces on wrist photoplethysmography prestudy for a wearable patient monitor. *Biomed Tech (Berl).* 2004;49(1–2):22–6.
28. Wong C, Zhang ZQ, Lo B, Yang GZ. Wearable sensing for solid biomechanics: a review. *IEEE Sens J.* 2015;15(5):2747–60.
29. Lee JM, Youngwon K, Welk GJ. Validity of consumer-based physical activity monitors. *Med Sci Sports Exerc.* 2014;46(9):1840–8.
30. Río-Rodríguez D, Iglesias-Soler E, Fernandez del Olmo M. Set configuration in resistance exercise: muscle fatigue and cardiovascular effects. *PLoS One.* 2016;11(3):e0151163.
31. Piirainen JM, Tanskanen M, Nissila J, et al. Effects of a heart rate-based recovery period on hormonal, neuromuscular, and aerobic performance responses during 7 weeks of strength training in men. *J Strength Cond Res.* 2011;25(8):2265–73.