

# Accuracy of Consumer Monitors for Estimating Energy Expenditure and Activity Type

JAMES A. WOODMAN<sup>1,2</sup>, SCOTT E. CROUTER<sup>1</sup>, DAVID R. BASSETT JR<sup>1</sup>, EUGENE C. FITZHUGH<sup>1</sup>, and WILLIAM R. BOYER<sup>1</sup>

<sup>1</sup>Department of Kinesiology, Recreation, and Sport Studies, The University of Tennessee Knoxville, Knoxville, TN; and <sup>2</sup>Department of Medicine, The University of Vermont, Burlington, VT

## ABSTRACT

WOODMAN, J. A., S. E. CROUTER, D. R. BASSETT JR, E. C. FITZHUGH, AND W. R. BOYER. Accuracy of Consumer Monitors for Estimating Energy Expenditure and Activity Type. *Med. Sci. Sports Exerc.*, Vol. 49, No. 2, pp. 371–377, 2017. **Introduction:** Increasing use of consumer-based physical activity (PA) monitors necessitates that they are validated against criterion measures. Thus, the purpose of this study was to examine the accuracy of three consumer-based PA monitors for estimating energy expenditure (EE) and PA type during simulated free-living activities. **Methods:** Twenty-eight participants (mean  $\pm$  SD: age, 25.5  $\pm$  3.7 yr; body mass index, 24.9  $\pm$  2.6 kg·m<sup>-2</sup>) completed 11 activities ranging from sedentary behaviors to vigorous intensities. Simultaneous measurements were made with an Oxycon portable calorimeter (criterion), a Basis Peak and Garmin Vivofit on the nondominant wrist, and three Withings Pulse devices (right hip, shirt collar, dominant wrist). Repeated-measures ANOVA were used to examine differences between measured and predicted EE. Intraclass correlation coefficients were calculated to determine reliability of EE predictions between Withings placements. Paired samples *t* tests were used to determine mean differences between observed minutes and Basis Peak predictions during walking, running, and cycling. **Results:** On average, the Basis Peak was within 8% of measured EE for the entire PA routine ( $P > 0.05$ ); however, there were large individual errors (95% prediction interval, -290.4 to +233.1 kcal). All other devices were significantly different from measured EE for the entire PA routine ( $P < 0.05$ ). For activity types, Basis Peak correctly identified  $\geq 92\%$  of actual minutes spent walking and running ( $P > 0.05$ ), and 40.4% and 0% of overground and stationary cycling minutes, respectively ( $P < 0.001$ ). **Conclusions:** The Basis Peak was the only device that did not significantly differ from measured EE; however, it also had the largest individual errors. Additionally, the Basis Peak accurately predicted minutes spent walking and running, but not cycling. **Key Words:** ACCELEROMETRY, PUBLIC HEALTH, OXYGEN CONSUMPTION, INDIRECT CALORIMETRY, VALIDATION, ACTIVITY RECOGNITION

Objective monitoring can be a useful method to track physical activity (PA) and has been shown to help motivate sedentary individuals to increase their PA (5). Consumer-based objective PA monitors are widely available, and the use of these devices is becoming more widespread in the general population and research (1,16). Based on a US national survey, 69% of respondents reported tracking a health indicator for themselves or others using paper tracking, activity trackers, or other methods (10). In addition, 60% of respondents reported tracking weight, diet, or exercise routine for health.

To date, numerous consumer-based monitors have been validated in laboratory settings, including the Fitbit (7,15,17,18,22,25,28), Nike FuelBand (15,25,26), Jawbone Up (3,9,15,17), and Misfit Shine (3,9,17). Previous research has shown mixed results for how accurately consumer-based devices estimate energy expenditure (EE) (9). For example, the Fitbit and Fitbit Ultra have been shown to be strongly correlated with measured EE during walking and jogging; intraclass correlation coefficients (ICC) ranged from 0.56 to 0.72 and from 0.81 to 0.87, respectively (18). In contrast, the basis B1 (first generation) had a weak correlation to measured EE during 69 min of structured activities ( $r = 0.136$ ; mean absolute percentage error [MAPE], 23.5%) (15).

With the introduction of new consumer-oriented PA monitors (14) and continual updates to the algorithms of existing devices, it is important for researchers to validate these devices. To our knowledge, the Basis Peak (second generation) has not been included in any published EE validation studies, the Garmin Vivofit has been included in two validation studies examining HR and step count predictions (8) and EE (2), and the Withings Pulse has been included in three validation studies for step count (13) and EE (9,17) prediction. The Basis Peak also has a function to

Address for correspondence: Scott E. Crouter, Ph.D., Department of Kinesiology, Recreation, and Sport Studies, 1914 Andy Holt Ave, Knoxville, TN 37996; E-mail: scrouter@utk.edu.

Submitted for publication May 2016.

Accepted for publication August 2016.

0195-9131/17/4902-0371/0

MEDICINE & SCIENCE IN SPORTS & EXERCISE®

Copyright © 2016 by the American College of Sports Medicine

DOI: 10.1249/MSS.0000000000001090

identify time spent walking, running, and cycling, although this function has not been previously validated. In addition, some of these devices can supposedly be worn on different parts of the body. For example, placement site (hip, neck lanyard, and pants and shirt pockets) of the Omron HJ-112 pedometer has been shown to not affect step counts, although the hip placement provided the least amount of random error (12). It is unclear how placement site might affect other outcomes such as estimates of EE. The Withings Pulse instructions state that the device can be worn anywhere but previous research has not examined the effects of placement on the output of this device.

Thus, the purpose of this study was to: 1) examine the accuracy of estimated EE from the Basis Peak, Garmin Vivofit, and Withings Pulse, compared with portable indirect calorimetry, during 11 different structured PA; 2) investigate the relationship of EE predictions among three placement sites for the Withings Pulse; and 3) validate the Basis Peak's activity identification function, which estimates time spent in walking, running, and cycling.

## METHODS

Twenty-eight participants were recruited from The University of Tennessee, Knoxville, and surrounding areas. Participants were screened for exclusion criteria using the Physical Activity Readiness Questionnaire. Exclusion criteria included currently pregnant, obesity (body mass index  $\geq 30 \text{ kg}\cdot\text{m}^{-2}$ ), orthopedic or musculoskeletal issues that would limit activity, or not being able to run on a treadmill for 5 min at  $134.1 \text{ m}\cdot\text{min}^{-1}$  and 0% incline. The treadmill run was used to ensure the participants could complete the vigorous intensity activities. Prior to participation, participants signed an informed consent form. This study was approved by the University of Tennessee Institutional Review Board.

Participants were asked to abstain from alcohol and vigorous exercise for 24-h before data collection, and abstain from eating and consuming caffeine for 4 h prior. Weight and height were measured in light clothing and no shoes, using a physician's scale and stadiometer, respectively. Participants were fitted with a Polar HR monitor, a Basis Peak and Garmin VivoFit on the nondominant wrist, three Withings Pulses (dominant wrist, shirt collar, and right hip), and an Oxycon portable calorimeter. Participants were then asked to complete a structured PA routine consisting of 11 activities that lasted approximately 90 min. The activities were intended to be simulations of free-living activities to replicate how the device would be used in real life, thus we choose to use overground walking, running, and cycling at self-selected speeds versus controlling for speed on a treadmill or cycle ergometer. Participants completed 10 min of supine rest and 5 min of the other 10 activities, with a minimum of 2 min of transition time between activities. Activities were completed in the following order from lowest to highest energy cost, except for activity 9 (seated rest) which was

used to provide a longer break between the running and cycling activities:

1. Supine rest
2. Computer usage in a seated position
3. Folding clothes in a seated position
4. Sweeping a floor
5. Treadmill walking at  $80.5 \text{ m}\cdot\text{min}^{-1}$  and 7% incline
6. Continuously ascending and descending stairs
7. Overground walking at a self-selected pace on a sidewalk, track, or in a gym
8. Overground running at a self-selected pace on a sidewalk, track, or in a gym
9. Seated rest
10. Overground cycling outside on a standard bicycle at a self-selected pace
11. Cycling on a Lode ergometer at 100 W

The Oxycon Mobile (CareFusion Corp, San Diego, CA) is a portable breath-by-breath indirect calorimeter that provides measures of oxygen consumption ( $\dot{V}\text{O}_2$ ) and carbon dioxide production ( $\dot{V}\text{CO}_2$ ). The device has two units measuring  $126 \times 96 \times 41 \text{ mm}$  each, and a total weight of 950 g (including backpack, battery, and mask). This device has been shown to be valid measuring  $\dot{V}\text{O}_2$  and  $\dot{V}\text{CO}_2$  compared with the Douglas Bag method (20) and a laboratory-based metabolic cart (19). Before each test, the device was calibrated, which consisted of ambient air sampling, volumetric calibration with a 3-L syringe, and gas calibration using a mixture of 15.93%  $\text{O}_2$  and 4.92%  $\text{CO}_2$ .

The Basis Peak (Basis Science, Inc., San Francisco, CA) is a small ( $3.6 \times 2.7 \text{ cm}$ , has a 27.3 cm), lightweight (44 g), wrist-worn activity monitor that is water resistant up to a pressure of 5 atmospheres (atm). It has a battery life of 2–3 d and is charged through a docking station connected to a computer. Sensors within this device include a triaxial accelerometer, two thermometers, an optical blood-flow sensor, and a galvanic skin response sensor. Data from these sensors are used to estimate HR, steps taken, and gross EE that are displayed on a touchscreen. Additionally, the sensors are used to identify how many minutes are spent in each of three activities: walking, running, and cycling. A participant profile was created with the MyBasis application using the investigator's smartphone. The same smartphone was used to edit the profile for each participant (gender, age, height, and weight), and the phone was then synchronized with the Basis device. All data are stored on company servers, which were accessed via smartphone and a computer-based web browser.

The Garmin VivoFit (Garmin Ltd., Schaffhausen, Switzerland) is a water-resistant (50 m), wrist-worn activity monitor that is small ( $2.1 \times 1.05 \text{ cm}$ ) and lightweight (25.5 g). It includes two band sizes to accommodate wrist circumferences ranging from 12 to 21 cm, and uses a coin cell battery that provides a battery life of up to 1 yr. A digital readout displays estimates of steps taken, ambulatory distance (i.e., walking and running distance), and gross EE. A participant

profile was created with the Garmin Connect application using the investigator's smartphone. The same smartphone was used to edit the profile for each participant (gender, age, height, and weight), and the phone was then synchronized to the Garmin device.

The Withings Pulse (Withings, Issy les Moulineaux, France) is a small ( $4.3 \times 2.2 \times 0.8$  cm), lightweight (8 g) device that is not water-resistant. It can estimate steps, ambulatory distance (i.e., walking and running distance), and net EE. This device does not require participant data to be entered before use, and according to the instructions, it can be worn on the hip, shirt collar, or either wrist using attachment straps provided by the manufacturer.

Breath-by-breath  $\dot{V}O_2$  and  $\dot{V}CO_2$  from the Oxycon were used to compute EE (kcal), which was averaged over a 15-s period, and used as the criterion measure for EE. EE data were analyzed for the entire PA routine and for each individual activity. Oxycon EE values were obtained for the entire routine (including transitions) by summing all 15-s values. For the 5-min individual activities, minutes 2:30 to 4:30 from the Oxycon were used to calculate the EE ( $\text{kcal} \cdot \text{min}^{-1}$ ) for each activity.

EE data for each consumer-based activity monitor are available via each manufacturer's smartphone or computer application; however, Withings and Garmin applications do not present minute-by-minute EE data. As such, EE values for all devices were recorded on a data sheet immediately before and after each activity, and the difference between the end EE and start EE for each activity was divided by the activity duration to get an estimated kilocalories per minute for each activity. Because the Withings Pulse provides estimates of net EE, whereas the Garmin VivoFit and Basis Peak estimate gross EE, we chose to convert all values to gross EE, so a direct comparison could be made. Thus, basal metabolic rate for each participant was calculated using the Harris-Benedict equation (11), which was added to the net EE value from the Withings Pulse to obtain estimates of gross EE.

The Basis Peak identified minutes spent in structured walking, running, and cycling that were obtained via the MyBasis app (<https://app.mybasis.com/>). The activity routine started on the minute according to the internal clock in the basis, such that the basis measures of time spent in structured activities could be compared with direct observation of these behaviors. All structured activity bouts were started on the minute, but not all bouts ended precisely on the minute. To ensure only valid data were included, the first and last whole minute of each activity bout were excluded from this analysis.

All analyses were conducted using IBM SPSS statistics software version 22 (IBM, Armonk, NY). For all analyses, an alpha of 0.05 was used to denote statistical significance, and data are presented as mean  $\pm$  SD, unless otherwise noted. Repeated-measures ANOVA were used to examine differences between measured (Oxycon) and estimated (consumer-based monitors) EE for the entire PA routine and each structured activity. When necessary, pairwise comparisons with Bonferroni adjustment were used to determine where significant differences existed. Modified Bland-Altman plots (4) were created to show the range of each monitor's individual error; dashed lines represent the 95% prediction interval (95% PI) and the solid line represents the mean error. The plots are modified by only including the measured values on the *x*-axis, which is acceptable when a "gold standard" is used for comparison. Accurate devices will have a narrow 95% PI and a mean error close to zero. MAPE (absolute values of the percent error relative to the Oxycon) was also calculated as an additional indicator of measurement error.

A repeated-measures ANOVA was used to test for mean differences between the three Withings Pulse placement sites during the entire activity routine and for each individual activity. When needed, pairwise comparisons with Bonferroni adjustments were used to find which placement sites were significantly different. ICC was calculated to examine reliability among the three Withings Pulse placement sites over the entire PA routine. Excellent, fair to good, and poor reliability were defined as ICC values of  $\geq 0.75$ , 0.4 to 0.75,  $< 0.4$ , respectively (21).

Paired samples *t* tests were used to determine mean differences between directly observed minutes and Basis Peak predicted minutes of treadmill walking, overground walking, overground running, overground cycling, and stationary cycling.

## RESULTS

Physical characteristics of the participants are presented in Table 1. One participant's Oxycon data file was not retrievable due to a download error. On some occasions, head movement during testing caused temporary occlusions in the Oxycon sampling line affecting eight of 297 individual activity bouts that were excluded from the analysis of the entire PA routine and individual activities: overground running (3), seated rest (2), overground cycling (1), and stationary cycling (2).

Table 2 shows the mean measured and estimated gross EE for all 11 individual activities and for the entire PA routine

TABLE 1. Physical characteristics of participants.

	Male ( <i>n</i> = 20)	Female ( <i>n</i> = 8)	All Participants ( <i>N</i> = 28)
Age (yr)	26.3 $\pm$ 4.9 (21.5–33.8)	23.4 $\pm$ 1.9 (21.5–26.3)	25.5 $\pm$ 3.7 (21.5–33.8)
Height (cm)	179.7 $\pm$ 5.2 (171.5–191.0)	165.9 $\pm$ 5.0 (157.5–173.0)	175.7 $\pm$ 8.1 (157.5–191.0)
Weight (kg)	83.3 $\pm$ 8.4 (65.6–96.2)	62.3 $\pm$ 4.5 (56.2–68.8)	77.3 $\pm$ 12.2 (56.2–96.2)
BMI ( $\text{kg} \cdot \text{m}^{-2}$ )	25.8 $\pm$ 2.2 (21.2–29.9)	22.7 $\pm$ 2.5 (19.6–26.1)	24.9 $\pm$ 2.6 (19.6–29.9)

Values are mean  $\pm$  SD (range).  
BMI, body mass index.

(all minutes and all transitions between activities included). All devices were significantly different from measured values for eight or more activities with mean differences, ranging from 0.1 to 8.5 kcal·min<sup>-1</sup>. Seated rest was the only activity for which all devices were not significantly different from measured EE ( $P > 0.05$ ).

For the entire PA routine, there were significant differences between the measured and estimated EE from all activity monitors ( $P < 0.05$ ), except for the Basis Peak ( $P = 0.257$ ; Table 2). On average, the Basis Peak estimated EE was 7% higher than the measured EE. The Garmin VivoFit and all three Withings placement sites significantly underestimated measured EE by 41.6% to 64.4% ( $P < 0.001$ ). Figure 1 shows the MAPE for each device tested. The basis had the lowest MAPE (27.2%), whereas the other devices have MAPE between 40.3% and 63.7%.

Figure 2 shows the Bland–Altman plots for the gross EE during the entire PA routine. The Basis Peak had the lowest mean error (−28.7 kcal); however, it had large individual errors (95% PI, −290.4 to +233.1 kcal). Other devices had greater mean errors (+169.8 to +262.7 kcal), but less individual error than the Basis Peak; 95% PI, +93.8 to +271.8 kcal (Garmin VivoFit), +142.7 to +382.6 kcal (Withings wrist), +59.8 to +286.2 kcal (Withings shirt collar), and +56.7 to +282.8 kcal (Withings hip). In addition, all devices had a systematic bias with significant correlations ( $P < 0.05$ ) between the difference score ( $y$  axis, measured minus estimated) and the measured value ( $x$  axis) ranging from  $R^2 = 0.151$  (basis) to  $R^2 = 0.965$  (Withings on the wrist).

For the entire routine, predicted mean gross EE for the Withings Pulse shirt collar and hip placements were both significantly higher than the wrist placement ( $P < 0.001$ ), but not significantly different from each other ( $P > 1.00$ ). However, all three placement sites significantly underestimated measured EE ( $P < 0.05$ ). The shirt collar and hip placements had fair to good reliability (ICC, 0.558;  $P < 0.05$ ), whereas the wrist and hip, and wrist and collar had poor reliability (ICC, 0.058 and 0.094, respectively;  $P < 0.05$ ). For seated computer use, seated rest, and stationary cycling, there were no significant differences between the three Withings Pulse placement locations ( $P > 0.05$ ). For all other individual activities, the shirt collar and hip placements were both

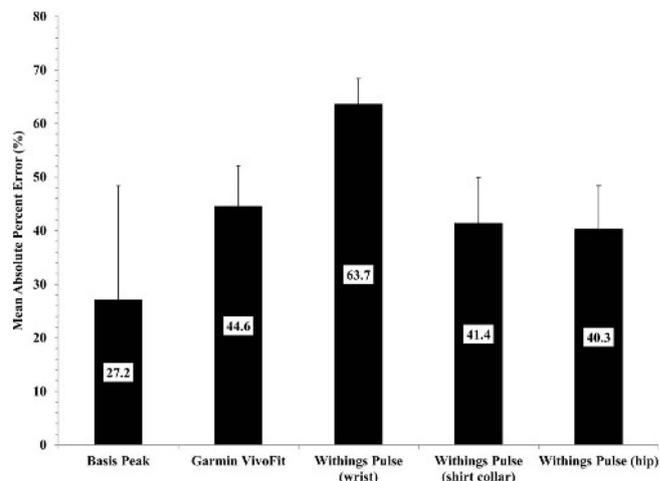


FIGURE 1—Mean absolute percent error ( $\pm$  SD) for estimation of gross energy expenditure.

significantly different from the wrist placement ( $P < 0.001$ ), but not significantly different from each other ( $P > 0.05$ ).

Compared with measured observed minutes, Basis Peak correctly predicted, on average, 92.9% 100%, and 94.7% of minutes spent during treadmill walking, overground walking, and overground running, respectively ( $P > 0.05$ ). However, it significantly underestimated the actual time spent in overground cycling (40.4% correctly identified;  $P < 0.001$ ) and stationary cycling (zero minutes correctly identified;  $P < 0.001$ ).

## DISCUSSION

The primary findings from the current study were that the second-generation Basis Peak predictions of gross EE were, on average, similar to measured EE over the entire structured PA routine; however, the Basis Peak EE estimates were significantly different from measured EE for eight activities. The current study also found that all three Withings Pulse placement sites and the Garmin VivoFit performed poorly and were significantly different from measured EE for the entire activity routine and most individual activities.

The current study found that the MAPE for the Basis Peak was 27.2%, whereas the other devices had MAPE ranging

TABLE 2. Mean  $\pm$  SD measured and estimated gross energy expenditure for each individual activity (kcal·min<sup>-1</sup>) and for the entire routine (kcal).

Activity	N	Oxycon	Basis Peak	Garmin VivoFit	Withings Wrist	Withings Shirt Collar	Withings Hip
Supine lying rest	27	1.5 $\pm$ 0.3	1.2 $\pm$ 0.3*	1.4 $\pm$ 0.3*	1.2 $\pm$ 0.2*	1.2 $\pm$ 0.2*	1.2 $\pm$ 0.2*
Seated computer use	27	1.8 $\pm$ 0.4	1.6 $\pm$ 0.6	1.7 $\pm$ 0.3	2.6 $\pm$ 0.4*	2.5 $\pm$ 0.3*	2.5 $\pm$ 0.3*
Seated folding towels	27	3.1 $\pm$ 0.5	2.7 $\pm$ 0.5*	2.0 $\pm$ 0.4*	2.7 $\pm$ 0.3*	2.5 $\pm$ 0.3*	2.5 $\pm$ 0.3*
Sweeping a floor	27	4.2 $\pm$ 0.8	2.8 $\pm$ 0.7*	2.5 $\pm$ 0.7*	2.8 $\pm$ 0.4*	2.5 $\pm$ 0.4*	2.6 $\pm$ 0.3*
Treadmill walking (80.5 m·min <sup>-1</sup> , 7% incline)	27	8.5 $\pm$ 1.6	10.1 $\pm$ 2.4*	4.2 $\pm$ 0.7*	3.4 $\pm$ 0.4*	5.6 $\pm$ 0.4*	5.7 $\pm$ 0.5*
Upstairs and downstairs	27	9.1 $\pm$ 1.3	10.3 $\pm$ 2.7*	4.1 $\pm$ 0.8*	3.9 $\pm$ 0.3*	7.2 $\pm$ 0.8*	7.3 $\pm$ 0.8*
Overground walking (average speed 79 m·min <sup>-1</sup> )	27	5.6 $\pm$ 1.1	10.3 $\pm$ 2.30*	4.2 $\pm$ 0.8*	3.5 $\pm$ 0.5*	5.7 $\pm$ 0.8	5.7 $\pm$ 0.8
Overground running (average speed 150 m·min <sup>-1</sup> )	24	13.7 $\pm$ 3.2	14.3 $\pm$ 2.6	10.8 $\pm$ 2.4*	5.2 $\pm$ 0.9*	13.5 $\pm$ 2.1	13.6 $\pm$ 2.0
Seated rest	25	2.5 $\pm$ 0.7	2.6 $\pm$ 3.3	1.7 $\pm$ 0.3	2.4 $\pm$ 0.2	2.5 $\pm$ 0.5	2.4 $\pm$ 0.5
Over-ground cycling (average speed 207 m·min <sup>-1</sup> )	26	8.7 $\pm$ 2.6	5.4 $\pm$ 3.5*	3.2 $\pm$ 7.1*	1.2 $\pm$ 0.2*	2.8 $\pm$ 0.7*	2.8 $\pm$ 0.8*
Stationary cycling (100 W)	25	9.2 $\pm$ 1.2	7.0 $\pm$ 5.1*	1.8 $\pm$ 0.6*	1.2 $\pm$ 0.2*	2.5 $\pm$ 0.4*	2.5 $\pm$ 0.4*
Entire PA routine	27	407.8 $\pm$ 71.4	436.5 $\pm$ 132.6	225.0 $\pm$ 43.1*	145.2 $\pm$ 17.4*	234.9 $\pm$ 31.6*	238.1 $\pm$ 31.3*

\*Significantly different from the Oxycon,  $P < 0.05$ .

from 40.3% to 63.7%. In addition, all devices had significant systematic biases with individual errors getting larger as the measured EE increased. Previous research has shown that the basis B1 (first generation) underestimated measured EE during a 69-min PA routine by 85.8 kcal (MAPE = 24%) and performed the worst compared with seven other devices. The best performers in that study were the Bodymedia Fit (MAPE = 9.3%) Fitbit Zip, (MAPE = 10.1%), and Fitbit One (MAPE = 10.4%) (15). In a separate study, the Withings Pulse (placed on the wrist) and Garmin Vivofit were shown to significantly underestimate total daily EE, measured by doubly labeled water, by 518 and 503 kcal·d<sup>-1</sup>,

respectively (17). This same study also showed that the Fitbit Flex was within 172 kcal·d<sup>-1</sup> of measured EE ( $P > 0.05$ ). Although it is difficult to draw comparisons between studies due to different criterion measures and methods used, there are some general trends emerging in the literature. The Fitbit is consistently among the best performers for estimating EE; however, we did not test a Fitbit device in the current study, thus we cannot draw conclusions on how it would compare with the devices used. A second common theme is that although some may have similar mean estimates of EE to measured values, all devices have large individual errors. This is an important consideration given in the current study

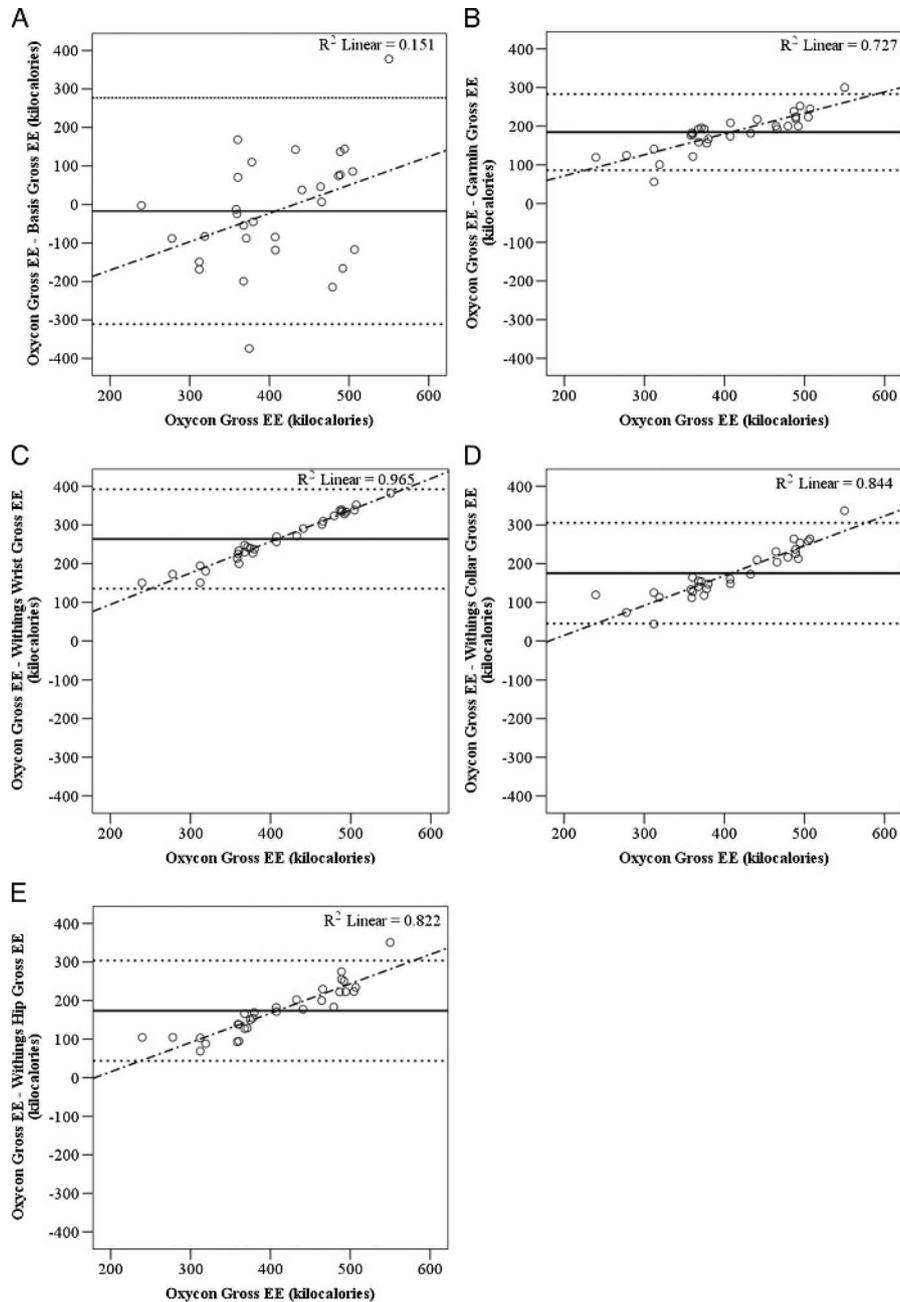


FIGURE 2—Bland–Altman plots of gross EE over the entire activity routine for (A) Basis Peak, (B) Gamin VivoFit, (C) Withings Pulse wrist, (D) Withings Pulse shirt collar, (E) Withings Pulse hip. *Solid line* represents the mean bias and *dashed horizontal lines* represent the 95% PI.

the Basis Peak was not significantly different from measured EE; however, the MAPE was 27.2%. Thus, on a group basis, the EE estimates were improved, but the MAPE was similar to previous research. A third consideration is that most devices tend to have larger individual errors at measured EEs as was seen in the current study and by Lee et al. (15).

Accurate estimates of daily EE are needed for individuals seeking weight loss through caloric restriction and/or increases in PA. Because these individuals may use consumer-based activity monitors for estimating their daily EE, the overestimations and underestimations seen during different activities may be less of a concern because they will potentially cancel each other out over the course of a day. However, if an individual performs most of their activities at the lower or higher end of the intensity spectrum, then this could impact the overestimation or underestimation resulting from these activity monitors. Based on the results of the current study, using the data from the entire routine, the Basis Peak has potential for providing valid total daily EE estimates on a group basis. However, caution is warranted when using the values at the individual level due to the large individual errors. Although the Withings Pulse and Garmin VivoFit had much lower individual error, their mean errors did not provide valid estimates of EE, thus should not be considered accurate on a group or individual level for estimating EE. Thus, consumers should be cautious when using these devices for the purpose of estimating daily EE, if energy balance is an important outcome. Another issue when using these devices is that most only provide a single EE value and obtaining minute-by-minute data can be challenging or impossible on most of these devices. Thus, users of these devices should be aware of the inherent limitations of how the devices work as obtaining EE data for a single workout session during the day may not be possible. The current study recorded EE values at the start and end of each activity bout as a way to overcome this issue; however, it is not ideal and those seeking EE data during a specific period should consider a similar approach if using devices with this limitation. It should be noted that some consumer activity trackers contain features that allow the user to “tag” or time structured exercise sessions and receive summative data for the specified period (e.g., Apple watch).

The Withings Pulse device provided consistent estimates of EE between the shirt collar and hip placement sites that were significantly higher EE estimates than the wrist location. However, all three locations were significantly different from measured EE. The device literature did not explicitly recommend any one placement site; sites were chosen based on the device accessories (wristband, clip), as well as images from the manufacturer’s instruction booklet and Website. Additionally, the device did not have an option to input the placement site, so the device does not have site-specific (e.g., wrist, shirt collar, hip) algorithms for predicting EE. Although it is interesting that the current study found that different wear locations for the same device provide similar results (shirt collar and hip), there is still a need for site-specific algorithms to predict EE, especially if the wrist location is being used.

The Basis Peak correctly identified more than 92% of the time spent walking and running, which could aid individuals in estimating total weekly walking and running time. More specifically, this information could help individuals in tracking activity to meet the current PA guidelines (27). It is important to note that the accuracy of this function is limited to walking and running; the device cannot identify stationary cycling at all, and only 40% of overground cycling minutes were identified. It is speculated that the Basis Peak uses the internal accelerometer to detect the different activity types. Walking and running are rhythmic activities and have been shown to be easily identifiable using accelerometers (6,23); however, cycling has always been problematic to detect with accelerometers, regardless of the placement location (23,24). With the wrist-worn Basis Peak, it is likely that the device is sensing bumps and vibrations that are transmitted through the bike to the wrist on the handlebars during over-ground cycling that are absent during stationary cycling. Nonetheless, the Basis Peak could still be used to encourage individuals who enjoy walking and running to increase their ambulatory activity.

Strengths of this study include the criterion measurements of EE (Oxycon) and directly observed activity type to validate the Basis Peak estimates. In addition, a wide range of activities were used, including typical activities of daily living. There were also some limitations to this study. The sample was a homogenous group of college-age adults, limiting generalizability to other populations. Future studies seeking to expand upon the validity of these consumer-based devices should consider a wider range of ages and fitness levels and extending the measurements to free-living environments.

## CONCLUSIONS

The primary finding from this study was that the Basis Peak was the only device tested that did not differ significantly from measured EE during the entire PA routine, showing potential for use in tracking daily EE; however, caution should be used due to the large individual errors for estimating EE across participants. All devices worked poorly for providing point estimates of EE during individual activities. Devices can also be affected by placement site and caution should be used when multiple wear locations are available without site-specific algorithms. Finally, the Basis Peak worked well for estimating time spent walking and running and showed promise for the prediction of individual activities in consumer devices, but it underestimated time spent in overground cycling by more than 50% and detected zero minutes of stationary cycling.

**Conflict of Interest:** The authors have no conflicts of interest. The results of the present study do not constitute endorsement by ACSM. The results of the study are presented clearly, honestly, and without fabrication, or inappropriate data manipulation.

## REFERENCES

- Albert MV, Deeny S, McCarthy C, Valentin J, Jayaraman A. Monitoring daily function in persons with transfemoral amputations using a commercial activity monitor: a feasibility study. *PM&R*. 2014;6(12):1120–7.
- Alsulbeeh SA, George AM, Baker A, Rohr LE, Basset FA. Accuracy of the vivofit activity tracker. *J Med Eng Technol*. 2016;40(6):298–306.
- Bai Y, Welk GJ, Nam YH, et al. Comparison of consumer and research monitors under semistructured settings. *Med Sci Sports Exerc*. 2016;48(1):151–8.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*. 1986;1(8476):307–10.
- Bravata DM, Smith-Spangler C, Sundaram V, et al. Using pedometers to increase physical activity and improve health: a systematic review. *JAMA*. 2007;298(19):2296–304.
- Crouter SE, Kuffel E, Haas JD, Frongillo EA, Bassett DR Jr. Refined two-regression model for the ActiGraph accelerometer. *Med Sci Sports Exerc*. 2010;42(5):1029–37.
- Diaz KM, Krupka DJ, Chang MJ, et al. Fitbit(R): an accurate and reliable device for wireless physical activity tracking. *Int J Cardiol*. 2015;185:138–40.
- El-Amrawy F, Nounou MI. Are currently available wearable devices for activity tracking and heart rate monitoring accurate, precise, and medically beneficial? *Health Inform Res*. 2015;21(4):315–20.
- Ferguson T, Rowlands AV, Olds T, Maher C. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. *Int J Behav Nutr Phys Act*. 2015;12(1):42.
- Fox S, Duggan M. Tracking for Health. Pew Research Center, Pew Internet and American Life Project 2013 [cited 2016 July 14]. Available from: <http://pewinternet.org/Reports/2013/Tracking-for-Health.aspx>.
- Harris JA, Benedict FG. A biometric study of human basal metabolism. *Proc Natl Acad Sci U S A*. 1918;4(12):370.
- Hasson R, Haller J, Pober D, Staudenmayer J, Freedson P. Validity of the Omron HJ-112 pedometer during treadmill walking. *Med Sci Sports Exerc*. 2009;41(4):805.
- Kooiman TJ, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers. *BMC Sports Sci Med Rehabil*. 2015;12(7):24.
- Lee J-M, Kim Y-W, Welk GJ. TRACK IT: Validity and utility of consumer-based physical activity monitors. *ACSMs Health Fit J*. 2014;18(4):16–21.
- Lee J-M, Kim Y, Welk GJ. Validity of consumer-based physical activity monitors. *Med Sci Sports Exerc*. 2014;46(9):1840–8.
- Meyer J, Hein A. Live long and prosper: potentials of low-cost consumer devices for the prevention of cardiovascular diseases. *Med 2 0*. 2013;2(2):e7.
- 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):702–3.
- Noah JA, Spierer DK, Gu J, Bronner S. Comparison of steps and energy expenditure assessment in adults of Fitbit Tracker and Ultra to the Actical and indirect calorimetry. *J Med Eng Technol*. 2013;37(7):456–62.
- Perret C, Mueller G. Validation of a new portable ergospirometric device (Oxycon Mobile) during exercise. *Int J Sports Med*. 2006;27(5):363–7.
- Rosdahl H, Gullstrand L, Salier-Eriksson J, Johansson P, Schantz P. Evaluation of the Oxycon Mobile metabolic system against the Douglas bag method. *Eur J Appl Physiol*. 2010;109(2):159–71.
- Rosner B. *Fundamentals of Biostatistics*. 7th ed. Belmont, CA: Cengage Learning; 2010. p. 569.
- Sasaki JE, Hickey A, Mavilia M, et al. Validation of the Fitbit wireless activity tracker for prediction of energy expenditure. *J Phys Act Health*. 2015;12(2):149–54.
- Staudenmayer J, Pober D, Crouter S, Bassett D, Freedson P. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. *J Appl Physiol*. 2009;107(4):1300–7.
- Steeves JA, Bowles HR, McClain JJ, et al. Ability of thigh-worn ActiGraph and activPAL monitors to classify posture and motion. *Med Sci Sports Exerc*. 2015;47(5):952–9.
- Storm FA, Heller BW, Mazzà C. Step detection and activity recognition accuracy of seven physical activity monitors. *PLoS One*. 2015;10(3):e0118723.
- Tucker WJ, Bhammar DM, Sawyer BJ, Buman MP, Gaesser GA. Validity and reliability of Nike + Fuelband for estimating physical activity energy expenditure. *BMC Sports Sci Med Rehabil*. 2015;7(1):14.
- U.S. Department of Health and Human Services. 2008 Physical Activity Guidelines for Americans Atlanta: U.S. Department of Health and Human Services; 2008 [cited 2016 April 15]. Available from: <http://www.health.gov/paguidelines>.
- 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.