Routledge
Taylor & Francis Group

Measurement in Physical Education and Exercise Science

ISSN: 1091-367X (Print) 1532-7841 (Online) Journal homepage: http://www.tandfonline.com/loi/hmpe20

Comparative Accuracy of a Wrist-Worn Activity Tracker and a Smart Shirt for Physical Activity Assessment

Alexander H. K. Montoye, Joe R. Mitrzyk & Monroe J. Molesky

MEASUREMENT IN PHYSICAL EDUCATION AND EXERCISE SCIENCE, 2017
https://doi.org/10.1080/1091367X.2017.1331166

Routledge
Taylor & Francis Group

Comparative Accuracy of a Wrist-Worn Activity Tracker and a Smart Shirt for Physical Activity Assessment

Alexander H. K. Montoye, Joe R. Mitrzyk, and Monroe J. Molesky

Department of Integrative Physiology and Health Science, Alma College, Alma, Michigan, USA

ABSTRACT

The purpose of the current study was to determine the accuracy of the Fitbit Charge HR and Hexoskin smart shirt. Participants (*n* = 32, age: 23.5 ± 1.3 years) wore a Fitbit and Hexoskin while performing 14 activities in a laboratory and on a track (lying, sitting, standing, walking various speeds and inclines, jogging, and cycling). Steps, kcals, heart rate, breathing rate, depth, and volume were measured by the Fitbit and Hexoskin and compared to criterion measures. The Fitbit and Hexoskin had low mean absolute percent error for steps (9.7%, 9.4%). The mean absolute percent error was low for heart rate (6.6% and 2.4%), with the Fitbit underestimating heart rate at higher intensities. Both devices had high mean absolute percent error for kcals (43.7% and 27.9%, respectively), and the Hexoskin had high mean absolute percent error for breathing rate, depth, and volume (19.4%, 35.6%, and 33.6%, respectively). The Fitbit and Hexoskin have utility for measurement of some, but not all, physical activity and physiologic variables which they measure.

KEYWORDS

smart clothing; textile; activity tracker; energy expenditure; consumer

Introduction

Regular participation in physical activity (PA) has long been known to provide substantial health benefits, including reduced risk of many chronic diseases and improved quality of life. However, studies in developed countries show that the majority of adults and children do not meet PA guidelines (Botey, Bayrampour, Carson, Vinturache, & Tough, 2016; Jefferis et al., 2014; Troiano et al., 2008). In order to determine adherence to guidelines, develop and monitor interventions to increase PA participation, and track changes in PA over time, accurate PA measurement tools are needed.

In recent years, interest in personal PA monitoring has increased dramatically with the use of smartphones and wearable devices (Almalki, Gray, & Sanchez, 2015; Evenson, Goto, & Furberg, 2015). These devices can provide real-time feedback on PA variables, such as steps taken, calories (kcals) burned, active time, and heart rate, and may be able to serve as motivational tools to get individuals more active (Cadmus-Bertram, Marcus, Patterson, Parker, & Morey, 2015; Chung, Skinner, Hasty, & Perrin, 2017; Hickey & Freedson, 2016; Naslund, Aschbrenner, & Bartels, 2016; Wang et al., 2015). However, these devices should be accurate to be most useful to the consumer. Underestimations of PA may result in discouragement or overtraining; conversely, overestimations of PA could result in failure to achieve PA recommendations or not being active enough to achieve appreciable health benefit.

A large number of devices are available on the market, with Fitbit monitors among the most popular activity tracking devices (Dolan, 2014). Evenson et al. (2015) conducted a systematic review of Fitbit and Jawbone activity monitors, two of the most popular consumer-based wearable device brands, and found that steps and kcals were the main variables assessed, with accuracy of step counting generally higher than that of kcal estimates. However, few studies have assessed the accuracy of wrist-worn, consumer-based activity monitors for assessment of heart rate, a variable of high importance for assessment of activity intensity. Until recently, the most common field method for assessing heart rate was via a chest strap, which would transfer data to a computer or wrist-worn watch. Wrist-based heart rate monitoring is a relatively new concept, where photoplethysmography is used to assess changes in light absorption due to the pulsatile flow of blood during the cardiac cycle. It is appealing in that it removes the need for a chest strap to assess heart rate. However, available findings for wrist-based heart rate suggest mixed results. The Fitbit Charge HR was among the first consumer-focused, wrist-based activity trackers on the market, and most of the limited available

CONTACT Alexander H. K. Montoye montoyeah@alma.edu Department of Integrative Physiology and Health Science, Alma College, 614 W. Superior, Alma, MI 48801

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hmpe.

research has focused on this device. A study by Wallen et al. (2016) found underestimations of heart rate predictions from four activity monitors (including Charge HR) by roughly 1%–9% overall in a mixture of activities. However, a brief report by Wang et al. (2017) found poor accuracy and wide limits of agreement for four monitors (including Charge HR) during walking at different speeds. Finally, a study by Jo et al. (2016) found that the Fitbit Charge HR had accurate heart rate assessment at lower intensity activities but that the Fitbit Charge HR had more pronounced underestimation once the heart rate was above 116 beats/minute. Given these limited and conflicting findings, more work is needed to elucidate how well, and for what types of activities, wrist-based heart rate assessment is suitable.

Clothing embedded with wearable sensors is a potential alternative to devices worn on the wrist or hip. Such "smart clothing," while not yet widely adopted, has potentially broad application to measurement of PA and physiologic variables during sport and everyday life. Clothing embedded with sensors removes the need for individuals to wear bulky and sometimes unsightly monitoring equipment, and it may improve compliance and quality of data collection. One such device, the Hexoskin, is a sleeveless compression shirt embedded with sensors to detect acceleration (for assessment of steps and kcals), heart rate, breathing rate, depth, and volume. The only known laboratory validation of the Hexoskin was conducted by Villar et al. (2015) and found low variability and good accuracy of the Hexoskin for assessment of heart rate, breathing rate, and breathing volume, but did not assess steps or kcal accuracy or compare Hexoskin results to other consumer-based wearables, limiting interpretation of their potential utility for PA measurement. Given the sparse literature evaluating heart rate accuracy from consumer-based wearables and lack of studies comparing wearable monitors to smart clothing, the purpose of this study was to validate and compare accuracy of the popular Fitbit Charge HR activity tracker to the Hexoskin biometric smart shirt for assessment of variables, including steps, kcals, and heart rate, for both devices, as well as breathing rate, depth, and volume for the Hexoskin during a variety of activities in a laboratory and track-based setting.

Methods

Participants

Healthy men (n = 18) and women (n = 14) aged 18–51 years and able to participate in moderate- or vigorous-intensity PA participated in this study. Participants were recruited via word-of-mouth, e-mail, and posted fliers, and provided informed consent approved by the Alma College Institutional Review Board prior to beginning the study.

Equipment

A Fitbit Charge HR (Fitbit Inc., San Francisco, CA, USA) activity tracker and Hexoskin (Carre Technologies Inc., Montreal, QC, Canada) biometric smart shirt were tested during this study. The Fitbit Charge HR is specifically designed for wear on the wrist and has a digital display which shows steps, kcals, and wrist-based heart rate (assessed using photoplethysmography); for this study the Fitbit was worn on participants' non-dominant wrist. The Charge HR was chosen because Fitbit is the most popular brand of consumer-based wearable monitors and the Charge HR was Fitbit's first available activity monitor which measured heart rate (Dolan, 2014). The Hexoskin is a compression shirt which is worn on the participants' torso and is connected to an included data logging device which worn on the right side of the torso (secured in a sewn-in pouch) and is connected via cable to sensors in the shirt. The Hexoskin data logger contains an accelerometer for assessment of steps and kcals and a global positioning system (GPS) unit for tracking of distance and position (when outside). The Hexoskin shirt contains sensors at the approximate level of the xiphoid process of the sternum and the umbilicus. These sensors are secured in place using elastic bands provided by the manufacturer and together assess single-lead electrocardiogram (ECG), (for assessment of heart rate) and strain gauges which allow for assessment of breathing rate, depth, and volume. The Hexoskin shirt comes in different sizes and is sex-specific; appropriate size and sex-specific shirts were used for each participant in this study.

The Fitbit and Hexoskin were initialized according to manufacturer specifications with the participant's sex, height, weight, and age at the beginning of the visit. Both the Fitbit and Hexoskin were synchronized to their associated applications on a smartphone to record steps, kcals, heart rate, breathing rate, depth, and volume (breathing variables for Hexoskin only). While the Hexoskin measured breathing depth (i.e., tidal volume), these data are not reported on the smartphone application, therefore, the breathing depth was calculated from the formula:

$$breathing\ rate\left(\frac{breaths}{minute}\right) \times breathing\ depth\left(\frac{L}{breath}\right) = breathing\ volume\left(\frac{L}{minute}\right)$$

(McArdle, Katch, & Katch, 2015). Algorithms for determination of PA and physiologic variables from both devices are proprietary.

An Omron HJ 323u Pedometer (Omron Corp., Osaka, Japan), Nonin PureSAT Pulse Oximeter (Nonin Medical Inc., Plymouth, MN, USA), and a Parvo TrueOne 2400 metabolic analyzer (Parvo Medics, Sandy, UT, USA) were used as criterion measures in this study. The Omron Pedometer HJ 323u, the criterion measure for steps, was placed by research staff over the participant's clothing on their right hip at the anterior axillary line. Step measurements were recorded before and after each activity throughout the visit. While this specific model of Omron has not been validated for measurement of steps, previous versions of the Omron HJ series (specifically the HJ 112, 113, 303, 151, and 720) have been shown to be highly accurate for step measurement during walking and running at various speeds compared to observer-counted steps (Giannakidou et al., 2012; Hasson, Haller, Pober, Staudenmayer, & Freedson, 2009; Holbrook, Barreira, & Kang, 2009; Steeves et al., 2011). For a randomly selected sample of five participants, steps were counted manually by a researcher to confirm the Omron pedometer as a valid measure of step counts for this study. The Nonin PureSAT Pulse Oximeter, the criterion for heart rate measurement, was placed on the participant's right middle finger to record heart rate during the last minute of each activity (in the case of the track-based activities, immediately following activity completion). The Nonin pulse oximeter has previously shown high accuracy compared to electrocardiogram-based heart rate (Gregoski et al., 2012). Finally, the Parvo metabolic analyzer served as the criterion measure for kcals, breathing rate, breathing depth, and breathing volume and was properly calibrated according to manufacturer specifications before each visit. The Parvomedics TrueOne 2400 system has previously been shown to provide highly accurate measures of oxygen consumption (hence, kcal measurement) and breathing volume compared to the gold-standard Douglas bag method (Bassett et al., 2001). Measurements were collected via a breathing mask worn by the participant while performing the laboratory activities only.

Protocol

Participants participated in 14 total exercises, 11 in the laboratory and three on a 200-meter indoor track. Laboratory exercises included lying down, sitting, standing, walking 2.0 miles/hour at 0.0% grade, walking 2.0 miles/hour at 5.0% grade, walking 2.0 miles/hour at 10.0% grade, walking 3.0 miles/hour at 0.0% grade, walking 3.0 miles/hour at 5.0% grade, brisk walking at 3.5–4.0 miles/hour, jogging at 5.0–10.0 miles/hour, and stationary cycling at 75–150 Watts. Walking and jogging exercises were performed on a treadmill (Mortara Trackmaster, Full Vision Inc., Newton, KS) controlled by research staff, and the stationary cycling exercise was performed on a stationary cycle (Monark 828 E, Monark Exercise AM, Vansbro, Sweden). Participants were allowed to choose the speed of the brisk walking, jogging, and cycling activities within the prescribed ranges, as long as the chosen speed was consistent for the entirety of the exercise. These exercises were performed for five minutes each, and the participant was allowed a break after each exercise. Track exercises included self-paced walking at both a leisure and brisk pace for 200 meters and self-paced jogging for 400 meters. Participants were allowed to self-pace themselves for all exercises performed on the indoor track as long as the speed was constant throughout the exercise. All participants were able to complete all exercises in the protocol. Length of the protocol (mean ± standard deviation) was 89.6 ± 5.5 minutes. Additionally, mean ± standard deviations were 3.6 ± 0.2 miles/hour for the brisk treadmill walk, 5.8 ± 0.9 miles/hour for the treadmill jog, 107.5 ± 13.9 Watts for the stationary cycling, 3.2 ± 0.4 miles/hour for the leisure track walk, 4.2 ± 0.5 miles/hour for the brisk track walk, and 7.4 ± 1.6 miles/hour for the track jogging. The Fitbit and Hexoskin were synchronized via Bluetooth to their associated applications and reported real-time data to researcher staff. Steps and kcals data were collected from the applications immediately before starting each exercise and immediately following the completion of each exercise; in this way, activity-specific measures of steps taken and kcals burned could be determined by subtracting the value immediately before the activity from the value immediately following the activity (e.g., if steps on the device was 651 before walking and 1,372 after walking, than the total steps taken during the walking activity was 1,372 – 651 = 721 steps). It took 1–2 minutes to sync the applications and record steps and kcals data for each activity; this time also served as a rest period for participants. Heart rate, breathing rate, breathing depth, and breathing volume data were recorded from the Hexoskin during the last minute of each exercise. On the indoor track, steps were recorded immediately before each exercise from the applications, and steps and heart rate were recorded immediately at the end of each exercise from the applications. In this way, data were collected for each activity, allowing for activity-specific analyses to be conducted to determine device accuracy. Kcals, breathing rate, breathing depth, and breathing volume were not recorded for the indoor track activities because the metabolic analyzer is not portable and could only be used in the laboratory.

Data cleaning and analysis

Prior to data analysis, data from the Fitbit and Hexoskin were screened for bad or missing data (which were excluded from analysis) using the following criteria: (1) if total kcals was recorded as 0 for an activity, it indicated that the device had not updated; (2) if heart rate, breathing rate, breathing depth, or breathing volume did not change for the Fitbit/Hexoskin during the activity, it indicated poor sensor connection.

For steps, the total number counted for each activity by the Omron, Fitbit, and Hexoskin were used in the analysis. For kcals, data recorded by the metabolic analyzer in minutes 4–5 of each activity were used to calculate a steady-state kcals/minute value, which was multiplied by five to compare with kcals recorded by the Fitbit and Hexoskin monitors during each activity. Heart rate data were recorded from the Fitbit and Hexoskin during the 5th minute of each activity and compared to steady-state heart rate assessed via the pulse oximeter. Finally, breathing rate, depth, and volume were recorded from the Hexoskin during the 5th minute of each activity and compared to steady-state values of each variable measured by the metabolic analyzer.

Overall and activity-specific predictions for the Fitbit and Hexoskin were compared to criterion-measured values for steps, kcals, and heart rate using repeated measures analysis of variance (within-subjects factor is device [criterion, Fitbit, Hexoskin]; no between-subjects factor used), with a Bonferroni correction to account for multiple comparisons; a total of 44 separate analyses were performed for totals, means, and activity-specific analyses for each variable (steps, kcals, heart rate). Additionally, paired t-tests were used to compare estimated breathing rate, breathing depth, and breathing volume from the Hexoskin to measured values of each of these variables by the criterion; in total, 36 paired t-tests were performed for this portion of the analysis. Mean absolute percent error (MAPE), a measure of the predictive accuracy of the devices which assesses differences between predicted and measured values without regard for under- or over-estimation, was calculated on an overall and activity-specific basis and for each variable; MAPE for variables assessed by both devices (steps, kcals, and heart rate) were compared between the Fitbit and Hexoskin using paired t-tests. Correlations between predicted and criterion-measured values were calculated for each variable and compared between the Fitbit and Hexoskin using paired t-tests. Finally, Bland-Altman plots were constructed to assess systematic overestimations or underestimations and limits of agreement by the devices. For all analyses, statistical significance was set at an alpha level of $p < 0.05$. Analyses were conducted in SPSS version 24.0 (IBM Corp, Armonk, NY) and Microsoft Excel (Microsoft Corp., Redmond, WA). In order to better interpret the accuracy of the devices, a MAPE threshold of <10% was used to characterize “low” measurement error, whereas ≥10% was considered “high” measurement error, as has been used in past work (Nelson, Kaminsky, Dickin, & Montoye, 2016).

Results

A total of 32 participants completed the study protocol; Table 1 displays their demographic information. The vast majority were right-hand dominant (93.8%). Data for a total of 15 activities (3.3% of data collected) were excluded due to poor synchronization of one or both the Fitbit and Hexoskin devices or poor device connection. In the subsample of five participants who had steps manually counted in addition to Omron-counted, MAPE for Omron-counted steps compared to manually counted steps was 10.7% across the three 2.0 miles/hour walking activities but was only 0.6% across the 3.0 miles/hour walks at 0% and 5% grades, 3.5–4.0 miles/hour walk at 0% grade, and both the leisure and brisk track walking activities and 0.5% for the treadmill and track jogging activities. Given the higher degree of error for step counting at the 2.0 miles/hour activities, data for steps are presented with all activities included and separately with the 2.0 miles/hour walking speeds excluded.

Table 2 displays overall and activity-specific means for steps, kcals, and heart rate for both the Fitbit and Hexoskin. Compared to the Omron, the Hexoskin significantly underestimated steps taken overall (8.7%–9.3%), also for all 2.0 miles/hour walking speeds (16.7%–44.0%), as well as the 3.0 miles/hour at 5% grade (5.9%), while the Fitbit underestimated steps overall (6.2%–6.5%) as well as for all level (0% grade) walking activities (4.7%–14.5%), walking at 3.0 miles/hour at 5% grade (8.8%), and the track-based leisure walking, brisk walking, and jogging

Table 1. Participant demographic information

	All (n = 32)	Male (n = 18)	Female (n = 14)
Age (years)	23.5 (1.3)	21.6 (1.1)	26.1 (2.6)
Height (in)	69.2 (0.6)	71.0 (0.7)	67.0 (0.9)
Weight (lbs)	174.9 (6.7)	190.7 (9.6)	154.6 (5.5)
Body mass index (kg/m^2)	25.6 (0.8)	26.6 (1.2)	24.2 (0.6)

Data are displayed as mean (standard error).

Table 2. Steps, kcals, and heart rate predicted by the Fitbit and Hexoskin compared to criterion measures

	Steps			Kcals			Heart rate (beats/minute)		
Activity	Criterion	Fitbit	Hexoskin	Criterion	Fitbit	Hexoskin	Criterion	Fitbit	Hexoskin
Totals	4785.5 (57.3)	4491.4 (62.8)*	4338.3 (71.2)*	271.9 (10.6)	344.9 (18.0)*	279.7 (19.9)	N/A	N/A	N/A
Means across all activities	344.0 (11.8)	321.8 (11.5)*	313.9 (11.6)*	18.5 (2.8)	17.9 (3.2)	20.3 (3.6)	104.9 (5.4)	101.6 (5.1)*	104.9 (5.6)
Means across all activities except 2.0 miles/hour walks	313.0 (14.5)	282.4 (13.9)*	307.1 (14.6)	N/A	N/A	N/A	N/A	N/A	N/A
Lie down	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	7.6 (0.3)	7.2 (0.3)	9.0 (1.3)*	63.5 (1.6)	62.5 (1.5)	62.0 (2.5)
Sit	0.0 (0.0)	0.2 (0.2)	0.0 (0.0)	7.1 (0.3)	7.2 (0.3)	7.5 (0.5)*	71.4 (1.9)	69.8 (1.7)	70.5 (2.9)
Stand	0.0 (0.0)	0.0 (0.0)	0.1 (0.0)	7.8 (0.4)	7.8 (0.4)	8.7 (0.8)*	81.2 (2.1)	77.9 (2.0)*	77.0 (3.7)
Walk—2.0 miles/hour 0% grade	471.1 (9.7)	449.0 (5.3)*	301.2 (18.0)*	17.1 (0.6)	17.9 (0.6)	33.3 (2.1)*	84.8 (1.8)	84.8 (2.1)	82.9 (3.2)
Walk—2.0 miles/hour 5% grade	451.7 (14.3)	479.4 (12.0)	338.5 (16.1)*	22.6 (0.8)	24.7 (0.8)	35.3 (1.9)*	95.0 (1.6)	95.8 (2.1)	96.7 (1.7)
Walk—2.0 miles/hour 10% grade	450.7 (10.6)	470.8 (9.8)	375.3 (11.5)*	29.9 (1.1)	32.9 (1.1)*	39.2 (2.0)*	107.4 (1.7)	106.3 (2.0)	107.5 (1.9)
Walk—3.0 miles/hour 0% grade	552.7 (4.3)	477.8 (20.9)*	542.3 (8.2)	22.3 (0.8)	23.4 (0.9)	39.5 (2.7)*	93.5 (1.9)	96.1 (2.2)	94.3 (2.0)
Walk—3.0 miles/hour 5% grade	542.9 (5.9)	495.2 (10.7)*	510.7 (13.1)*	30.3 (1.1)	33.1 (1.2)*	40.5 (2.9)*	107.2 (1.9)	105.4 (2.2)	107.6 (1.7)
Walk—3.5–4.0 miles/hour 0% grade	591.9 (5.9)	506.2 (15.2)*	597.2 (7.0)	26.3 (0.9)	28.3 (1.0)	42.1 (3.2)	102.1 (1.9)	100.4 (1.9)	102.4 (2.0)
Jog—5.0–10.0 miles/hour 0% grade	799.3 (8.8)	792.5 (8.2)	802.5 (9.4)	53.8 (2.1)	60.7 (2.4)*	59.4 (2.7)*	152.8 (2.2)	149.2 (2.4)*	153.2 (2.5)
Cycle—75–150 W	N/A	59.7 (8.2)	27.2 (9.4)	34.5 (1.1)	37.4 (1.2)*	41.3 (2.8)*	129.7 (3.4)	121.8 (5.0)*	130.9 (3.3)
Track walk—leisure pace	280.7 (8.1)	245.9 (8.1)*	280.6 (4.9)	N/A	N/A	N/A	101.5 (2.6)	100.9 (3.0)	102.0 (2.4)
Track walk—brisk pace	243.5 (7.2)	191.7 (7.2)*	246.2 (4.0)	N/A	N/A	N/A	116.4 (3.1)	103.2 (4.2)*	116.1 (3.0)
Track jog—self-selected pace	346.4 (11.6)	337.1 (11.6)*	342.6 (12.1)	N/A	N/A	N/A	166.3 (2.7)	151.6 (3.4)*	164.8 (2.7)

Data are displayed as mean (standard error).
N/A: Not applicable.
*Indicates significant difference from criterion ($p < 0.05$).
^Indicates non-significant trend toward significant difference from criterion ($0.05 \leq p < 0.10$).

activities (2.6%–21.3%). Conversely, both devices were not different from the criterion during sedentary/stationary activities (lying down, sitting, and standing). For kcals, both devices significantly overestimated kcals burned for the highest intensity activities (cycling and jogging; 8.4%–12.8% for Fitbit, 10.4%–19.7% for Hexoskin). Moreover, the Fitbit overestimated kcals burned overall (26.8%) and for two of the inclined walking activities (10.0% for 2.0 miles/hour 10% grade and 9.2% for 3.0 miles/hour 5% grade), while the Hexoskin overestimated kcals during all activities (5.6%–94.7%) except for the 3.5–4.0 miles/hour treadmill walk. Finally, the Hexoskin was not significantly different from the criterion measure for heart rate measured for any of the activities in the lab or on the track, whereas the Fitbit significantly underestimated heart rate overall (3.1%) and during standing, jogging, and cycling in the lab (2.4%–6.1%) and the brisk walk and jog on the track (8.9%–11.7%).

Table 3 displays breathing rate, depth, and volume measured by the metabolic analyzer compared to the predictions by the Hexoskin. The Hexoskin significantly overestimated breathing rate overall (12.2%), as well as at the 3.0 miles/hour walking at 5% grade (10.3%) and trended toward overestimation for 3.0 miles/hour walking at 0% grade, 3.5–4.0 miles/hour walking, and jogging. Breathing depth showed a trend toward being underestimated for standing and was significantly underestimated overall (20.7%) and for all other activities (16.3%–29.1%) except sitting. Consequently, the Hexoskin significantly underestimated breathing volume overall (17.5%) and for all activities (14.5%–30.3%) except for cycling, 2.0 miles/hour walk at 0% grade, and sitting.

Table 4 shows both overall and activity-specific MAPE for step, kcal, and heart rate predictions by the Fitbit and Hexoskin. Overall, both the Fitbit and Hexoskin had less than 10% error for steps and heart rate predictions, whereas kcal estimations by both devices had MAPE considerably higher than 10% (27.9%–43.7% overall and 12.5%–92.0% for activity-specific analyses). Additionally, MAPE for the Hexoskin was significantly lower than the Fitbit overall for both kcals and heart rate but for steps only when the 2.0 miles/hour walking activities were excluded. In activity-specific analyses, both devices had low MAPE (0.0%–6.5%) when estimating steps for lying down, sitting and standing. The Hexoskin had significantly higher MAPE than the Fitbit for each of the 2.0 miles/hour walking activities but lower MAPE for steps predictions for the 3.0 miles/hour walks at 0% and 5% grades, 3.5–4.0 miles/hour walk at 0% grade, and both the leisure and brisk track walking activities. Both devices had low MAPE for steps predictions during

Table 3. Overall means and MAPE for Hexoskin-predicted breathing rate, depth, and volume predictions compared to criterion measures

	Breathing rate (breaths/minute)		Breathing depth (L/breath)		Breathing volume (L/minute)		MAPE for breathing variables		
Activity	Criterion	Hexoskin	Criterion	Hexoskin	Criterion	Hexoskin	Breathing rate	Breathing depth	Breathing volume
Overall	23.8 (1.3)	26.7 (2.9)*	1.11 (0.04)	0.88 (0.06)*	28.0 (3.0)	23.1 (3.1)*	19.4 (4.3)	35.6 (4.7)	33.6 (3.8)
Lie down	16.0 (0.9)	15.6 (0.8)	0.64 (0.03)	0.46 (0.04)*	9.7 (0.4)	6.7 (0.4)*	8.7 (1.4)	32.2 (2.7)	35.5 (2.7)
Sit	15.7 (0.8)	14.8 (1.0)	0.64 (0.02)	0.69 (0.08)	9.8 (0.5)	9.0 (0.8)	15.4 (2.8)	45.2 (9.2)	40.2 (7.8)
Stand	16.2 (0.8)	18.1 (2.4)	0.68 (0.03)	0.56 (0.07)^	10.8 (0.5)	8.2 (0.7)*	24.3 (11.7)	44.0 (6.7)	42.1 (5.6)
Walk—2.0 miles/hour 0% grade	22.5 (0.8)	22.4 (1.7)	0.86 (0.03)	0.72 (0.06)*	18.9 (0.7)	17.0 (2.0)	15.2 (4.6)	32.4 (4.7)	41.6 (10.1)
Walk—2.0 miles/hour 5% grade	23.1 (0.8)	30.1 (6.7)	1.12 (0.04)	0.84 (0.07)*	25.4 (0.8)	18.8 (1.2)*	40.5 (31.0)	37.5 (4.2)	31.5 (3.9)
Walk—2.0 miles/hour 10% grade	24.5 (0.9)	26.3 (1.8)	1.37 (0.05)	0.96 (0.07)*	32.6 (1.0)	23.9 (1.6)*	11.9 (4.8)	35.0 (3.7)	33.1 (3.6)
Walk—3.0 miles/hour 0% grade	25.6 (0.8)	31.6 (3.6)^	1.03 (0.04)	0.80 (0.07)*	25.9 (0.9)	21.4 (1.3)*	27.7 (10.9)	35.6 (4.4)	29.5 (3.5)
Walk—3.0 miles/hour 5% grade	26.2 (0.9)	28.9 (1.8)*	1.31 (0.05)	0.98 (0.07)*	33.6 (1.2)	26.1 (1.5)*	13.1 (4.3)	35.0 (3.8)	30.4 (3.7)
Walk—3.5–4.0 miles/hour 0% grade	26.9 (0.9)	31.7 (2.8)^	1.13 (0.04)	0.90 (0.07)*	29.8 (1.1)	25.1 (1.6)*	28.3 (8.7)	33.1 (4.5)	26.7 (3.4)
Jog—5.0–10.0 miles/hour 0% grade	35.2 (1.3)	38.1 (2.0)^	1.89 (0.10)	1.51 (0.09)*	65.4 (2.8)	55.9 (3.5)*	17.4 (4.1)	26.1 (3.6)	24.9 (3.4)
Cycle—75–150 W	27.1 (1.2)	29.3 (1.4)	1.59 (0.07)	1.26 (0.12)*	42.8 (2.1)	37.6 (4.5)	11.8 (2.9)	35.7 (4.7)	39.7 (7.9)

Data shown as mean (standard error).
MAPE: Mean absolute percent error.
*Indicates significant difference from criterion ($p < 0.05$).
^Indicates non-significant trend toward significant difference from criterion ($0.05 \le p < 0.10$).

Table 4. Overall and activity-specific MAPE for Fitbit and Hexoskin steps, kcals, and heart rate predictions

	Steps		Kcals		Heart rate (beats/minute)	
Activity	Fitbit	Hexoskin	Fitbit	Hexoskin	Fitbit	Hexoskin
Overall	9.7 (1.2)	9.4 (0.7)	43.7 (3.4)	27.9 (2.7)*	6.6 (0.6)	2.7 (0.3)*
Overall (not including 2.0 miles/hour walks)	8.7 (1.0)	3.8 (0.6)*	N/A	N/A	N/A	N/A
Lie down	0.0 (0.0)	0.0 (0.0)	58.8 (12.4)	18.8 (2.1)*	4.7 (0.9)	3.8 (0.6)
Sit	3.1 (3.1)	0.0 (0.0)	25.5 (2.9)	36.2 (8.8)	6.5 (0.7)	3.4 (0.9)*
Stand	0.0 (0.0)	6.5 (4.5)	27.8 (7.0)	81.4 (12.7)*	5.6 (0.8)	5.3 (2.3)
Walk—2.0 miles/hour 0% grade	7.6 (1.5)	36.9 (3.3)	92.0 (9.9)	23.9 (3.9)*	7.9 (1.8)	3.0 (0.5)*
Walk—2.0 miles/hour 5% grade	17.2 (4.8)	28.8 (3.1)*	47.9 (4.8)	18.9 (2.8)*	5.5 (1.3)	3.3 (1.1)
Walk—2.0 miles/hour 10% grade	13.4 (4.6)	21.0 (3.2)*	22.9 (4.2)	19.3 (2.8)	4.3 (0.8)	2.7 (0.5)
Walk—3.0 miles/hour 0% grade	14.8 (3.7)	4.7 (0.9)*	74.2 (6.6)	24.4 (3.9)*	6.0 (1.5)	2.3 (0.6)*
Walk—3.0 miles/hour 5% grade	9.8 (1.8)	9.2 (1.9)*	26.5 (7.2)	17.8 (2.9)	3.7 (1.1)	2.0 (0.5)
Walk—3.5–4.0 miles/hour 0% grade	15.0 (2.5)	1.9 (0.5)	50.5 (11.3)	21.5 (3.2)*	4.6 (1.0)	2.0 (0.4)*
Jog—5.0–10.0 miles/hour 0% grade	1.3 (0.2)	1.2 (0.6)*	12.5 (1.6)	18.7 (2.7)	3.3 (0.7)	1.9 (0.6)
Cycle—75–150 W	N/A	N/A	32.2 (3.8)	31.5 (4.6)	9.2 (2.4)	2.1 (0.6)*
Track walk—leisure pace	13.4 (2.4)	4.2 (1.0)*	N/A	N/A	5.7 (1.1)	2.1 (0.4)*
Track walk—brisk pace	23.1 (2.6)	2.6 (0.8)*	N/A	N/A	14.9 (2.8)	2.1 (0.7)*
Track jog—self-selected pace	4.4 (0.8)	4.0 (1.9)	N/A	N/A	10.3 (2.0)	2.1 (0.5)*

Data are displayed as mean (standard error).
MAPE: Mean absolute percent error.
N/A: Not applicable.
*Indicates significant difference from Fitbit ($p < 0.05$).

both laboratory and track jogging activities. For kcals, the Hexoskin had significantly lower MAPE than the Fitbit for lying down, standing, walking 2.0 miles/hour at 0% and 5% grades, and walking 3.0 and 3.5–4.0 miles/hour at 0% grade. However, neither device achieved the <10% MAPE for any activity. Conversely, for heart rate, both devices had MAPE <10% for all activities in the lab; the Hexoskin also had <10% MAPE for all track activities and had significantly lower MAPE than the Fitbit for sitting, all 0% grade walking speeds (2.0, 3.0, and 3.5–4.0 miles/hour), and cycling as well as all three track activities. Activity-specific MAPE for breathing rate, depth, and volume are shown in Table 3. Lying down had a MAPE <10% for breathing rate; otherwise, MAPE was >10% overall and for all other activities for breathing rate, depth, and volume.

Table 5 displays correlations for the Fitbit and Hexoskin compared to criterion-measured values for each variable. The Fitbit and Hexoskin had equally high correlations with measured steps (r = 0.94 versus 0.95, respectively; p = 0.39). Conversely, the Fitbit had significantly lower correlations with criterion-measured values than the Hexoskin for kcals (0.84 versus 0.94, p = 0.001) and heart rate (0.90 versus 0.98, p = 0.001). Additionally, the Hexoskin had moderately strong ($r \geq 0.60$) or strong ($r \geq 0.80$) correlations with criterion measures for all breathing variables (Safrit & Wood, 1995).

Figures 1–6 present Bland-Altman plots for measured versus predicted values for each of the variables assessed. For steps, kcals, and heart rate, the Hexoskin had narrower 95% limits of agreement (indicating better overall agreement) than the Fitbit. For steps, the Hexoskin had the poorest predictive accuracy in the 350–450 step range (2.0 miles/hour walking activities, cadence ~70–85 steps/min), underestimating steps taken compared to the criterion. Conversely, the Fitbit showed the poorest step prediction accuracy in the 400–500 step range, without consistent overestimation or underestimation. For both the Fitbit and Hexoskin, limits of agreement were wide for kcal prediction, with poorer predictive accuracy (predominately overestimation) occurring for higher intensity activities. The Fitbit also showed poorer accuracy for assessment of heart rate primarily during higher intensity activities (underestimation), with wide overall 95% limits of agreement (–17.4 to +21.2 beats/min). Conversely, the Hexoskin had much narrower 95% limits of agreement (–8.2 to +8.8 beats/min) and similar accuracy across intensities. Finally, the Hexoskin showed overestimation of breathing rate at higher intensities, while there was no apparent relationship between intensity and predictive error for breathing depth or volume.

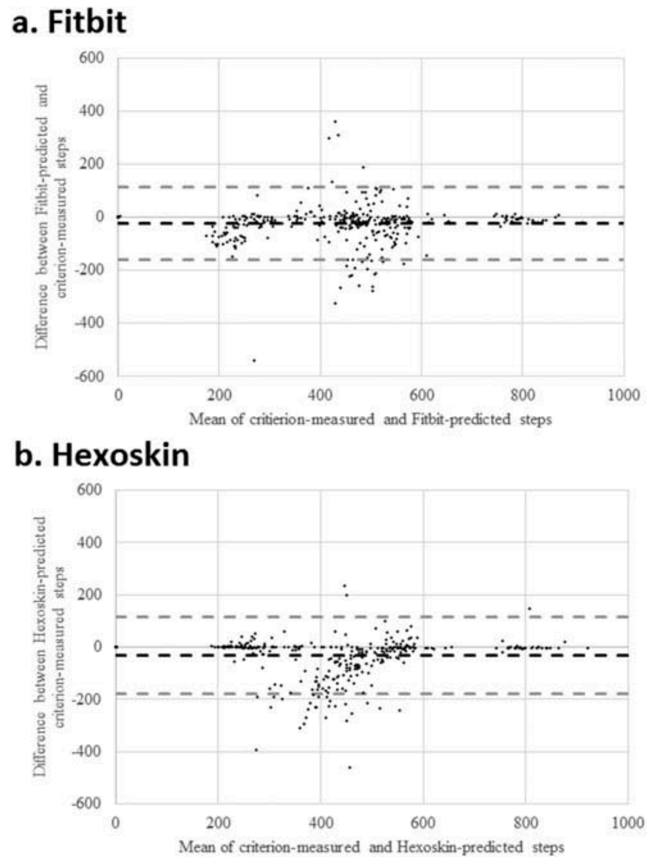

Figure 1. Bland-Altman plots for measured and predicted steps. Grey dashed lines represent ±1.96 standard deviations from mean. Black dashed line represents mean.

Table 5. Correlations between Fitbit and Hexoskin predictions with criterion measurements

	Fitbit	Hexoskin
Steps	0.94 (0.01)	0.95 (0.01)
Kcals	0.84 (0.02)	0.94 (0.01)*
Heart rate	0.90 (0.02)	0.98 (0.01)*
Breathing rate	N/A	0.79 (0.05)
Breathing depth	N/A	0.80 (0.04)
Breathing volume	N/A	0.92 (0.02)

Data are shown as mean (standard error).
N/A: Not applicable.
*Indicates significant difference from Fitbit (p < 0.05).

Discussion

This study's purpose was to compare the accuracy of the popular wrist-worn Fitbit Charge HR activity tracker to a Hexoskin biometric smart shirt for assessment of steps, kcals, and heart rate, as well as assess the accuracy of the Hexoskin for breathing rate, depth, and volume measures. Using the somewhat arbitrary threshold of <10% for MAPE, both devices had low overall measurement error for assessing steps taken and heart rate, with the Hexoskin having significantly lower MAPE than the Fitbit for

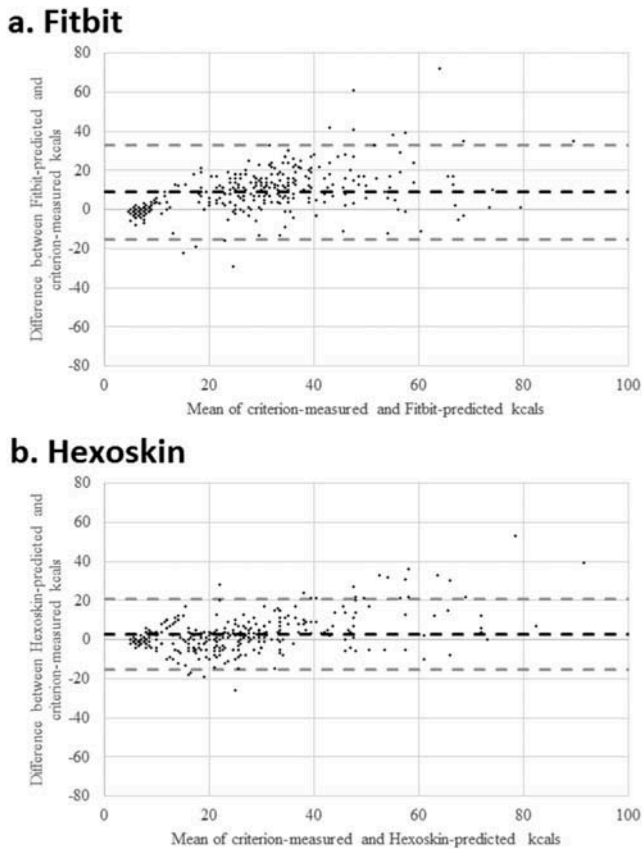

Figure 2. Bland-Altman plots for measured and predicted kcals.
Grey dashed lines represent ±1.96 standard deviations from mean. Black dashed line represents mean.

heart rate and, when 2.0 miles/hour walks were excluded, steps. Conversely, both devices had high measurement error for kcals (although MAPE for the Hexoskin was significantly lower than the Fitbit) and the Hexoskin had high measurement error for breathing rate, depth, and volume.

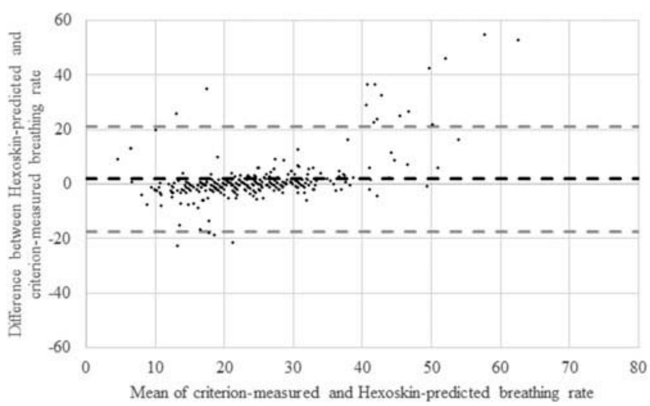

Figure 4. Bland-Altman plot for measured and Hexoskin-predicted breathing rate.
Grey dashed lines represent ±1.96 standard deviations from mean. Black dashed line represents mean. Breathing rates displayed in breaths/minute.

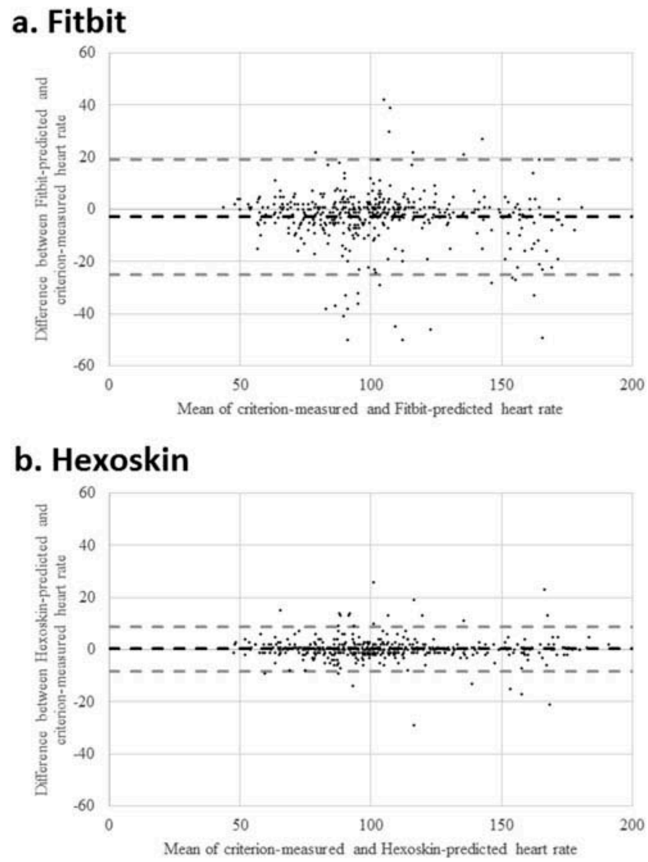

Figure 3. Bland-Altman plots for measured and predicted heart rate.
Grey dashed lines represent ±1.96 standard deviations from mean. Black dashed line represents mean. Heart rate displayed as beats/min.

In terms of steps, the Hexoskin had poor prediction accuracy for slow walking speeds when compared to the Omron, although our subanalysis of five participants found ~11% error for step counting by the Omron for the 2.0 miles/hour walks. The poorer

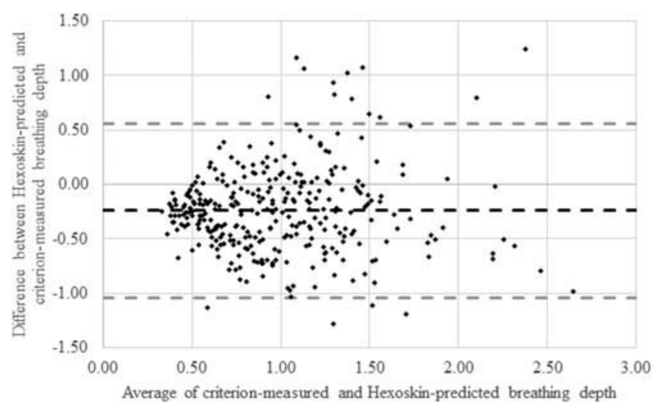

Figure 5. Bland-Altman plot for measured and Hexoskin-predicted breathing volume.
Grey dashed lines represent ±1.96 standard deviations from mean. Black dashed line represents mean. Breathing depths displayed in L/breath.

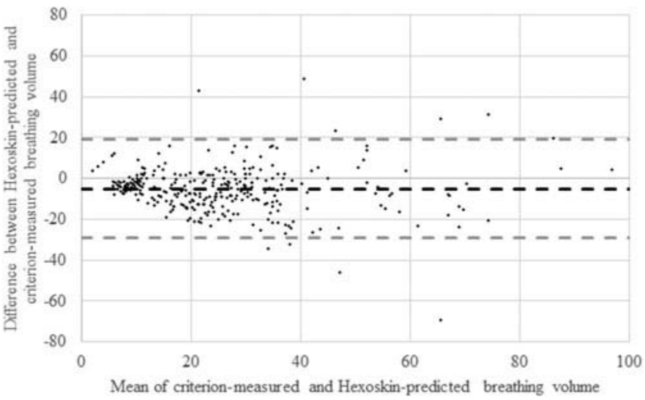

Figure 6. Bland-Altman plot for measured and Hexoskin-predicted breathing volume.

Grey dashed lines represent ±1.96 standard deviations from mean. Black dashed line represents mean. Breathing volumes displayed in L/minute.

accuracy of both the Hexoskin and Omron for slow walking speeds is common with activity monitors worn close to the center of mass (Ehrler, Weber, & Lovis, 2016; Giannakidou et al., 2012). However, the Omron performed with <1% MAPE compared to manually counted steps for all other activities, rendering it a useful criterion for stationary activities as well as walking and jogging ≥3.0 miles/hour. While five participants is too small of a sample to confidently declare the accuracy of the Omron, we did conduct a secondary analysis with the 2.0 miles/hour walks excluded to enhance confidence in the findings for the rest of the activities, which showed that the Hexoskin performed well at tracking steps during activities of 3.0 miles/hour or higher. Conversely, the Fitbit had less consistency in its measurement error for steps (i.e., no detectable pattern between certain activity types/speeds and measurement errors), making it less predictable in terms of how it would perform when used outside the laboratory.

For heart rate, it is notable that the Hexoskin performed with low MAPE for activities of all intensities, both in the laboratory and on the track. Our study is in agreement with the previous Hexoskin validation, which also showed high accuracy for heart rate assessment and no systematic or activity-specific bias (Villar et al., 2015). The high accuracy of the Hexoskin for heart rate measures is not surprising given that the heart rate sensors in the shirt are similarly positioned to chest-based heart rate straps, which have been shown to be highly accurate in previous work (Laukkanen & Virtanen, 1998; Wang et al., 2016). The Fitbit also had low MAPE for heart rate measurement for low-intensity activities but higher MAPE for cycling as well as activities on the track (i.e., outside the laboratory), which resulted in a ~ 3% underestimation across the protocol. A study by Wallen et al. (2016) evaluating accuracy of the Fitbit Charge HR for heart rate accuracy found similar results, with a ~ 9% underestimation overall and more pronounced underestimation at higher intensities. Jo et al. (2016) further support this finding, showing that heart rate underestimation was more pronounced for activities above a heart rate of 116 beats/minute. Another study by Wang et al. (2017) examining the Fitbit Charge HR found a slightly lower correlation with measured heart rate than our study (0.84 versus 0.90) as well as slightly wider 95% limits of agreement (−24 to +39 versus −17 to +21 beats/minute). Other wrist-based heart rate devices show similar measurement errors, indicating that it is more likely the current limitations of the technology rather than a specific brand of device (Spierer, Rosen, Litman, & Fujii, 2015; Wang et al., 2016). Thus, the available evidence suggests that the Fitbit and other wrist-based heart rate monitors cannot match the accuracy of heart rate monitors assessed via chest strap. For individuals looking to track resting heart rates or heart rates during low-intensity activities, the accuracy of the Fitbit and other wrist-based wearable monitors may be acceptable; however, individuals looking to use the devices to track heart rate during higher intensity PA should be aware of likely underestimation by the Fitbit Charge HR and consider chest-based heart rate monitors for optimal accuracy.

Despite strong correlations with measured kcals, predicted kcals from both the Hexoskin and Fitbit were poor, with overall MAPE ranging ~28–48% and poor activity-specific kcal predictions. Companies producing consumer-focused activity monitors typically do not reveal the algorithms used to take raw data and convert it to meaningful outcomes (i.e., steps, kcals, etc.). However, it is likely that these devices use a combination of steps or acceleration data along with heart rate to determine kcals burned. Thus, the higher kcal prediction accuracy of the Hexoskin compared to the Fitbit is likely at least partly attributed to the Hexoskin's higher heart rate accuracy. Other studies evaluating activity-specific kcal estimations of wearable activity monitors also point to poor predictive accuracy for individual activities performed for a short duration (Nelson et al., 2016; Sasaki et al., 2014). However, analyses of longer time periods (e.g., a whole protocol versus single activities or entire days) show lower predictive errors because overestimations from certain activities are washed out by underestimations for other activities (Lee, Kim, & Welk, 2014; Murakami et al., 2016). This washout effect, while lowering overall error, is still problematic as it makes it difficult to assess day-to-day differences and does not given an indication for whom or what types of activities the device will have the highest or lowest predictive error. Due to the high MAPE values found for kcal predictions in this study, as well as in a previous work, we recommend that kcal estimates from both the Fitbit Charge HR and Hexoskin be interpreted cautiously.

Finally, our study found poor prediction of breathing rate, depth, and volume from the Hexoskin. While breathing rate had the lowest MAPE and was not different from the criterion for most activities in the activity-specific analysis, breathing depth was underestimated for most activities, which contributed to underestimation of breathing volume for most activities. Given similar MAPE across activities for all breathing variables, inaccuracy appears equally spread across activities. Additionally, other than potential overestimation of breathing rate when breathing rate is high, no other trends were present in the data. Our findings stand in contrast to those of Villar et al., who found high accuracy of breathing rate and volume (e.g., correlations of 0.96–1.00 with measured breathing rate and mean differences of <0.5 breaths/minute) across a variety of activities and intensities (Villar et al., 2015). The reason for differing findings between studies is unclear. Both studies used well-supported criterion measures of respiration, a variety of activities, and participants with similar demographics. Since the respiration and heart rate sensors are embedded in the shirt, it is conceivable that the sensors could get damaged with repeated use or washing of the shirts. However, this was unlikely a contributing factor since our lab has several older (1–2 years old) and several brand new shirts, but the data show no differences in accuracy by shirt age. Additionally, heart rate accuracy was consistently high in our study, minimizing the likelihood of damage to the shirts' sensors. While we followed manufacturer directions and used appropriately sized, sex-specific shirts in the study, it may be that the shirts may fit some individuals better than others. Depending on the sensitivity of the strain gauges within the shirt, even small deviations from the optimal fit may affect accuracy of the sensors for measuring the breathing variables. Importantly, the correlations between measured and predicted values for each of the breathing variables were moderately strong or strong, meaning that the Hexoskin could be used to assess relative changes in breathing rate, depth, and volume across different activities. However, the actual values for each variable should be interpreted cautiously.

Our study has several notable strengths and limitations. Strengths include (1) the number and variety of activities performed in the laboratory and also on the track; (2) high-quality criterion measures; (3) direct comparability of two types of measurement devices; and (4) assessment of several variables from each device on an overall and activity-specific basis. Study limitations include (1) a relatively homogenous sample in age, with many of the participants being recreationally active or involved in a club or varsity sport; (2) few sport activities evaluated, which may be important to individuals considering using these devices; (3) no assessment of device accuracy in a free-living setting, which does not always mirror findings in a laboratory setting; and (4) a sample size sufficient to detect medium or large effect sizes (i.e., ≥0.5) (Cohen, 1988) but which was underpowered to detect smaller effect sizes, which may be of importance in some contexts. These limitations should be considered and addressed in future research.

In conclusion, our study found high accuracy of the Hexoskin biometric smart shirt for assessment of steps during all activities except slow walking and high accuracy of heart rate across all activities, but lower accuracy for kcals, breathing rate, breathing depth, and breathing volume. The Fitbit Charge HR achieved high accuracy for steps taken and for heart rate but tended to underestimate heart rates for higher-intensity activities and had lower accuracy for assessing kcals. These results should be taken into consideration when determining if these devices are appropriate for use in specific sport- or health-related contexts.

Acknowledgments

There are no conflicts of interest to report. Dr. John Davis assisted with data interpretation and study findings.

Funding

This study was partially funded by a grant from the Herbert H. and Grace A. Dow Foundation.

References

Almalki, M., Gray, K., & Sanchez, F. M. (**2015**). The use of self-quantification systems for personal health information: Big data management activities and prospects. *Health Information Science and Systems*, *3*(Suppl 1), S1. doi:10.1186/2047-2501-3-S1-S1

Bassett, Jr., D. R., Howley, E. T., Thompson, D. L., King, G. A., Strath, S. J., McLaughlin, J. E., & Parr, B. B. (**2001**). Validity of inspiratory and expiratory methods of measuring gas exchange with a computerized system. *Journal of Applied Physiology*, *91*(1), 218–224.

Botey, P., Bayrampour, H., Carson, V., Vinturache, A., & Tough, S. (**2016**). Adherence to Canadian physical activity and sedentary behaviour guidelines among children 2 to 13 years of age. *Preventive Medicine Reports*, *3*, 14–20. doi:10.1016/j.pmedr.2015.11.012

Cadmus-Bertram, L. A., Marcus, B. H., Patterson, R. E., Parker, B. A., & Morey, B. L. (**2015**). Randomized trial of a Fitbit-based physical activity intervention for women. *American Journal of Preventive Medicine*, *49*(3), 414–418. doi:10.1016/j.amepre.2015.01.020

Chung, A. E., Skinner, A. C., Hasty, S. E., & Perrin, E. M. (**2017**). Tweeting to health: A novel mHealth intervention using Fitbits and Twitter to foster healthy lifestyles. *Clinical Pediatrics*. doi:10.1177/0009922816653385

Cohen, S. (**1988**). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Dolan, B. (**2014**). *Fitbit, Jawbone, Nike had 97 Percent of Fitness Tracker Retail Sales in 2013*. Retrieved from http://mobihealthnews.com/28825/fitbit-jawbone-nike-had-97-percent-of-fitness-tracker-retail-sales-in-2013/

Ehrler, F., Weber, C., & Lovis, C. (**2016**). Influence of pedometer position on pedometer accuracy at various walking speeds: A comparative study. *Journal of Medical Internet Research*, *18*(10), e268. doi:10.2196/jmir.5916

Evenson, K. R., Goto, M. M., & Furberg, R. D. (**2015**). Systematic review of the validity and reliability of consumer-wearable activity trackers. *The International Journal of Behavioral Nutrition and Physical Activity*, *12*, 159. doi:10.1186/s12966-015-0314-1

Giannakidou, D. M., Kambas, A., Ageloussis, N., Fatouros, I., Christoforidis, C., Venetsanou, F., … Taxildaris, K. (**2012**). The validity of two Omron pedometers during treadmill walking is speed dependent. *European Journal of Applied Physiology*, *112*(1), 49–57. doi:10.1007/s00421-011-1951-y

Gregoski, M. J., Mueller, M., Vertegel, A., Shaporev, A., Jackson, B. B., Frenzel, R. M., … Treiber, F. A. (**2012**). Development and validation of a smartphone heart rate acquisition application for health promotion and wellness telehealth applications. *International Journal of Telemedicine and Applications*, *2012*, 1–7. doi:10.1155/2012/696324

Hasson, R. E., Haller, J., Pober, D. M., Staudenmayer, J., & Freedson, P. S. (**2009**). Validity of the Omron HJ-112 pedometer during treadmill walking. *Medicine and Science in Sports and Exercise*, *41*(4), 805–809. doi:10.1249/MSS.0b013e31818d9fc2

Hickey, A. M., & Freedson, P. S. (**2016**). Utility of consumer physical activity trackers as an intervention tool in cardiovascular disease prevention and treatment. *Progress in Cardiovascular Diseases*, *58*(6), 613–619. doi:10.1016/j.pcad.2016.02.006

Holbrook, E. A., Barreira, T. V., & Kang, M. (**2009**). Validity and reliability of Omron pedometers for prescribed and self-paced walking. *Medicine and Science in Sports and Exercise*, *41*(3), 670–674. doi:10.1249/MSS.0b013e3181886095

Jefferis, B. J., Sartini, C., Lee, I. M., Choi, M., Amuzu, A., Gutierrez, C., … Whincup, P. H. (**2014**). Adherence to physical activity guidelines in older adults, using objectively measured physical activity in a population-based study. *BMC Public Health*, *14*, 382. doi:10.1186/1471-2458-14-382

Jo, E., Lewis, K., Directo, D., Kim, M. J., & Dolezal, B. A. (**2016**). Validation of biofeedback wearables for photoplethysmographic heart rate tracking. *Journal of Sports Science & Medicine*, *15*(3), 540–547.

Laukkanen, R. M., & Virtanen, P. K. (**1998**). Heart rate monitors: State of the art. *Journal of Sports Sciences*, *16* (Suppl), S3–7. doi:10.1080/026404198366920

Lee, J. M., Kim, Y., & Welk, G. J. (**2014**). Validity of consumer-based physical activity monitors. *Medicine and Science in Sports and Exercise*, *46*(9), 1840–1848. doi:10.1249/MSS.0000000000000287

McArdle, W. D., Katch, F. I., & Katch, V. L. (**2015**). Exercise physiology: Nutrition, energy, and human performance (8th ed.). Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins.

Murakami, H., Kawakami, R., Nakae, S., Nakata, Y., Ishikawa-Takata, K., Tanaka, S., & Miyachi, M. (**2016**). Accuracy of wearable devices for estimating total energy expenditure: Comparison with metabolic chamber and doubly labeled water method. *JAMA Internal Medicine*, *176*(5), 702–703. doi:10.1001/jamainternmed.2016.0152

Naslund, J. A., Aschbrenner, K. A., & Bartels, S. J. (**2016**). Wearable devices and smartphones for activity tracking among people with serious mental illness. *Mental Health and Physical Activity*, *10*, 10–17. doi:10.1016/j.mhpa.2016.02.001

Nelson, M. B., Kaminsky, L. A., Dickin, D. C., & Montoye, A. H. (**2016**). Validity of consumer-based physical activity monitors for specific activity types. *Medicine and Science in Sports and Exercise*, *48*(8), 1619–1628. doi:10.1249/MSS.0000000000000933

Safrit, M., & Wood, T. (**1995**). *Introduction to measurement in physical education and exercise science* (3rd ed.). St. Louis, MO: Mosby.

Sasaki, J. E., Hickey, A., Mavilia, M., Tedesco, J., John, D., Kozey Keadle, S., & Freedson, P. S. (**2014**). Validation of the Fitbit wireless activity tracker for prediction of energy expenditure. *Journal of Physical Activity & Health*, *12*(2), 149–154. doi:10.1123/jpah.2012-0495

Spierer, D. K., Rosen, Z., Litman, L. L., & Fujii, K. (**2015**). Validation of photoplethysmography as a method to detect heart rate during rest and exercise. *Journal of Medical Engineering & Technology*, *39*(5), 264–271. doi:10.3109/03091902.2015.1047536

Steeves, J. A., Tyo, B. M., Connolly, C. P., Gregory, D. A., Stark, N. A., & Bassett, D. R. (**2011**). Validity and reliability of the Omron HJ-303 tri-axial accelerometer-based pedometer. *Journal of Physical Activity & Health*, *8*(7), 1014–1020. doi:10.1123/jpah.8.7.1014

Troiano, R. P., Berrigan, D., Dodd, K. W., Masse, L. C., Tilert, T., & McDowell, M. (**2008**). Physical activity in the United States measured by accelerometer. *Medicine and Science in Sports and Exercise*, *40*(1), 181–188. doi:10.1249/mss.0b013e31815a51b3

Villar, R., Beltrame, T., & Hughson, R. L. (**2015**). Validation of the Hexoskin wearable vest during lying, sitting, standing, and walking activities. *Applied Physiology, Nutrition, and Metabolism*, *40*(10), 1019–1024. doi:10.1139/apnm-2015-0140

Wallen, M. P., Gomersall, S. R., Keating, S. E., Wisloff, U., & Coombes, J. S. (**2016**). Accuracy of heart rate watches: Implications for weight management. *PLoS One*, *11*(5), e0154420. doi:10.1371/journal.pone.0154420

Wang, J. B., Cadmus-Bertram, L. A., Natarajan, L., White, M. M., Madanat, H., Nichols, J. F., … Pierce, J. P. (**2015**). Wearable Sensor/device (Fitbit One) and SMS text-messaging prompts to increase physical activity in overweight and obese adults: A randomized controlled trial. *Telemedicine and E-Health*, *21* (10), 782–792. doi:10.1089/tmj.2014.0176

Wang, J. B., Cataldo, J. K., Ayala, G. X., Natarajan, L., Cadmus-Bertram, L. A., White, M. M., … Pierce, J. P. (**2016**). Mobile and wearable device features that matter in promoting physical activity. *Journal of Mobile Technology in Medicine*, *5*(2), 2–11. doi:10.7309/jmtm.5.2.2

Wang, R., Blackburn, G., Desai, M., Phelan, D., Gillinov, L., Houghtaling, P., & Gillinov, M. (**2017**). Accuracy of wrist-worn heart rate monitors. *JAMA Cardiology*, *2*(1), 104–106. doi:10.1001/jamacardio.2016.3340