

Original Paper

# Automatic Identification of Physical Activity Type and Duration by Wearable Activity Trackers: A Validation Study

Diana Dorn<sup>1</sup>, MS; Jessica Gorzelitz<sup>1</sup>, MS; Ronald Gangnon<sup>2</sup>, PhD; David Bell<sup>1</sup>, PhD; Kelli Koltyn<sup>1</sup>, PhD; Lisa Cadmus-Bertram<sup>1</sup>, PhD

<sup>1</sup>Department of Kinesiology, University of Wisconsin-Madison, Madison, WI, United States

<sup>2</sup>Department of Population Health Sciences, University of Wisconsin-Madison, Madison, WI, United States

**Corresponding Author:**

Lisa Cadmus-Bertram, PhD

Department of Kinesiology

University of Wisconsin - Madison

2000 Observatory Drive

Madison, WI, 53562

United States

Phone: 1 608 265 5946

Email: [cadmusbertra@wisc.edu](mailto:cadmusbertra@wisc.edu)

## Abstract

**Background:** Activity trackers are now ubiquitous in certain populations, with potential applications for health promotion and monitoring and chronic disease management. Understanding the accuracy of this technology is critical to the appropriate and productive use of wearables in health research. Although other peer-reviewed validations have examined other features (eg, steps and heart rate), no published studies to date have addressed the accuracy of automatic activity type detection and duration accuracy in wearable trackers.

**Objective:** The aim of this study was to examine the ability of 4 commercially available wearable activity trackers (Fitbits Flex 2, Fitbit Alta HR, Fitbit Charge 2, and Garmin Vivosmart HR), in a controlled setting, to correctly and automatically identify the type and duration of the physical activity being performed.

**Methods:** A total of 8 activity types, including walking and running (on both a treadmill and outdoors), a run embedded in walking bouts, elliptical use, outdoor biking, and pool lap swimming, were tested by 28 to 34 healthy adult participants (69 total participants who participated in some to all activity types). Actual activity type and duration were recorded by study personnel and compared with tracker data using descriptive statistics and mean absolute percent error (MAPE).

**Results:** The proportion of trials in which the activity type was correctly identified was 93% to 97% (depending on the tracker) for treadmill walking, 93% to 100% for treadmill running, 36% to 62% for treadmill running when preceded and followed by a walk, 97% to 100% for outdoor walking, 100% for outdoor running, 3% to 97% for using an elliptical, 44% to 97% for biking, and 87.5% for swimming. When activities were correctly identified, the MAPE of the detected duration versus the actual activity duration was between 7% and 7.9% for treadmill walking, 8.7% and 144.8% for treadmill running, 23.6% and 28.9% for treadmill running when preceded and followed by a walk, 4.9% and 11.8% for outdoor walking, 5.6% and 9.6% for outdoor running, 9.7% and 13% for using an elliptical, 9.5% and 17.7% for biking, and was 26.9% for swimming.

**Conclusions:** In a controlled setting, wearable activity trackers provide accurate recognition of the type of some common physical activities, especially outdoor walking and running and walking on a treadmill. The accuracy of measurement of activity duration varied considerably by activity type and tracker model and was poor for complex sets of activity, such as a run embedded within 2 walking segments.

(*JMIR Mhealth Uhealth* 2019;7(5):e13547) doi: [10.2196/13547](https://doi.org/10.2196/13547)

**KEYWORDS**

fitness trackers; exercise; accelerometry; data accuracy

## Introduction

### Background

Adequate physical activity participation is one of the most important behaviors people can adopt to maintain their health and well-being. Physical activity reduces the risk of several major chronic diseases [1] and early mortality [2], reduces health risks associated with overweight and obesity [3], and improves psychological outcomes, including mood and energy [4]. For adults, the 2018 American federal aerobic physical activity guidelines prescribe  $\geq 150$  min per week of moderate-intensity activity or  $\geq 75$  min of vigorous-intensity physical activity or an equivalent combination of the two [5]. Results from US national surveillance estimates show suboptimal rates of activity nationwide and highlight the importance of promoting physical activity participation to increase the overall health of the population [6].

Consumer-grade activity trackers are one tool that may help individuals increase and monitor their physical activity participation. These devices are available to consumers at a relatively low cost, with approximately 14 million Fitbits alone sold in 2018 [7] and 120 million devices projected to be sold by 2019 [8]. Such trackers have been shown to support increased physical activity participation in adults [9,10] and are suitable for incorporation into clinical research and health promotion interventions [9,11,12]. Naturally, the extent of the utility of these devices depends on the accuracy of their features. Peer-reviewed independent research is critical to understanding and validating these technologies, particularly as manufacturers typically neither provide comprehensive results of internal studies nor release their proprietary classification algorithms. Previous validation studies have reported high correlations between device step counts and the criterion [13,14] and a general underestimation of energy expenditure as compared with criterion measurements [15], and heart rate validation studies have shown that wearable devices are more accurate during rest than during moderate exercise [16-18].

With the evolution of physical activity algorithmic-based classification, manufacturers are developing beyond-the-manual logging of activities (eg, a user tagging an activity as *running* or *cycling*) toward use of pattern recognition algorithms to automatically detect certain activity types. Automatic activity recognition—called SmartTrack on Fitbit trackers [19] and Move IQ on Garmin trackers [20]—allows a wearable activity tracker to recognize and classify specific activity types, without input from the user. If accurate, this feature could enhance health-based research by lending insight into what types of

activities participants are performing and what could aid researchers in tracking adherence to physical activity guidelines and fidelity to interventions.

### Objectives

Given the importance of understanding the accuracy of trackers' new automatic activity detection capabilities, and the lack of current validation studies available on this feature, the objective of this study was to explore the accuracy of this feature on 4 commercially available wearable activity trackers in a controlled setting: the Fitbit Flex 2, the Fitbit Charge 2, the Fitbit Alta HR (Fitbit), and the Garmin Vívosmart HR (Garmin). Specifically, this study aimed to determine (1) the accuracy of the activity-type identification and (2) the accuracy of the measured duration of activity in a controlled setting.

## Methods

### Participants

This study protocol was approved by the University of Wisconsin-Madison Health Sciences Institutional Review Board. Written informed consent was obtained for all participants. Participants were adults aged from 18 to 50 years, who were able to perform moderate- and vigorous-intensity physical activity and were free from specific health limitations, including heart attack, angioplasty or heart surgery in the previous 3 months, active chest pain, shortness of breath, fainting, angina pectoris, current pregnancy, or any physical disability that would preclude use of a treadmill, an elliptical, or a bicycle. Participants were excluded from swimming if they reported an inability to swim front crawl (ie, freestyle) for 15 min. Participants were also excluded if their measured resting systolic blood pressure (via Riester ri-champion N automatic cuff, Riester) was  $> 180$  mmHg or resting diastolic blood pressure was  $> 110$  mmHg.

### Study Design

The study comprised 4 activity modules (Table 1) designed to reflect some of the most common types of moderate-to-vigorous intensity physical activity with the intention of evaluating 34 trials per module on the basis of previous validation studies with small sample sizes [15]. Power analyses were not conducted because of a lack of previously published (or *a priori*) identified effect sizes. Participants self-selected which module or modules they wished to complete. Participants could complete more than 1 module if desired, but each module was scheduled on a different day to avoid excessively long activity participation in a single day by each participant. Participation in multiple modules was not expected to influence results.

**Table 1.** Description and duration of activities in each of the experimental modules.

Module and activities	Activity duration
<b>Module A<sup>a</sup></b>	
Treadmill walk (2-4.5 mph)	15 min
Treadmill run (>4.5 mph)	15 min
Embedded run (walk, run, walk) <sup>b</sup>	25 min (5-min walk, 15-min run, 5-min walk)
<b>Module B<sup>a</sup></b>	
Outdoor walk	15 min
Outdoor run	15 min
Elliptical	15 min
<b>Module C</b>	
Bike	15 min
<b>Module D</b>	
Swim	15 min

<sup>a</sup>Activities in this module separated by 10 min of rest.

<sup>b</sup>These activities were completed at the same miles per hour (mph) designation as the previous activities. The purpose of the embedded run was to test the trackers' ability to detect a run with walking before and after.

## Activity Modules

*Activity Module A* was completed on a treadmill (Trackmaster and Precor Inc) and comprised 15 min of walking (at speeds of 2 to 4.5 mph), 10 min of stationary rest, 15 min of running (>4.5 mph), 10 min of stationary rest, and a 15-min run embedded in 5-min segments of walking to simulate a warm up and cool down before and after a run (walk-run-walk). The purpose of the treadmill run embedded in walking bouts was to test the trackers' ability to detect the running bout when preceded and followed by bouts of walking. This activity is referred to as the *embedded run*. *Activity Module B* comprised walking outdoors on a continuous path for 15 min, 10 min of stationary rest, running on a path for 15 min, 10 min of stationary rest, and using an elliptical trainer with moving arm handles (Precor Inc) for 15 min. Participants were instructed to use the moving arm handles on the elliptical machine for the activity's duration. *Activity Module C* comprised 15 min of outdoor cycling on a continuous path (referred to as *bike* or *biking*). *Activity Module D* comprised 15 min of swimming in a 25-yard, 8-lane indoor pool. Participants began wearing devices following initialization in the laboratory; then they walked to the treadmill, pool, outdoor walking path, or other activity location (<5 min). Following each activity within the module, participants began the bout of sedentary rest. No sedentary rest was observed after the last activity in each module. Directly following each activity, participants reported a rating of perceived exertion (RPE) using the Borg RPE scale, which has high validity and reliability for reporting relative exercise intensity [21,22], to assess the intensity at which the trackers might successfully detect an activity and duration. Participants removed the trackers immediately after the last activity in each module.

## Devices

The 4 devices used in this study were the Fitbit Flex 2, the Fitbit Charge 2, the Fitbit Alta HR (Fitbit), and the Garmin Vivosmart

HR (Garmin). These devices were chosen because at the time of the study, they were the only ones that offered the automatic activity detection feature. Participants wore all 4 devices concurrently, 2 on each wrist, and placement was randomized—although previous validation work [16] demonstrated that tracker accuracy does not differ by wrist location. For Module D, only the Fitbit Flex 2 was worn as the others do not detect swimming. Following informed consent, participants completed a brief health history form and a demographics questionnaire, and height and weight were measured once using a stadiometer (Health o meter, Welch Allyn). All trackers were initialized using their respective Fitbit or Garmin apps, using each participant's actual height, weight, sex, and age. The same 4 devices were reinitialized and used for each participant in the study to minimize the potential risk of interdevice heterogeneity. The Fitbit devices were programmed to detect activities lasting longer than 10 min (the shortest amount of time required for activity-type recognition by the devices).

## Outcome Measures

Each activity performed was observed and recorded in conjunction with a timer on a laptop, which was used to record the duration of the activity; this was considered the criterion measure. The timer was stopped when the participant came to a complete stop at the end of the activity, and the duration was recorded. Upon completion of each module, the trackers were synced to their respective apps on a laptop, and the trackers' designation of activity type and duration were recorded from the apps. A correct identification occurred when the tracker's activity designation matched the actual activity performed. An erroneous identification occurred when the tracker's activity designation did not match the actual activity performed, and a missed identification occurred when the tracker did not designate any activity when there was an activity performed.

## Statistical Analysis

All statistical analyses were completed using Statistical Analysis Software version 9.4 (SAS Institute Inc.) and Microsoft Excel version 1811 (Microsoft). Data from all users were inspected for errors before analysis. One participant's elliptical trainer data were excluded because of failure to follow protocol (the participant did not use the moving arm handles). A total of 2 participants who were unable to complete the swimming activity were excluded from analyses. Descriptive statistics were used to characterize the participant sample and the mean Borg RPE of each activity. Frequency counts were calculated to analyze the number of correct identifications and the amount of missed and erroneous identifications. Summary statistics and box plots were produced to analyze the duration of activities that were correctly identified by the trackers. Mean absolute percent error (MAPE) ( $(1/n * | \text{true duration} - \text{device-based duration measure} |)$ ) was calculated for each activity and device to establish the differences between the criterion duration (duration measured with a laptop timer) and device measured duration.

## Results

Participants in this study (N=69) were 61% female, had a mean age of 26.4 (SD 8.7) years, and had an average body mass index of 23.9 (SD 4.1) kg/m<sup>2</sup>. Of the 69 participants, 36 (52%) participants chose to complete 1 module, 10 (15%) participants chose to complete 2 modules, 16 (23%) participants chose to complete 3 modules, and 7 (10%) participants chose to complete all 4 modules. Although the study was designed for 34 trials per module, the actual number of trials used in the analysis varied slightly because of infrequent but occasional inability to sync devices and retrieve data between study visits. The mean Borg RPE rating was 8.6 (SD 1.8) for the treadmill walk (corresponds to an *extremely light* to *very light* intensity); 13.6 (SD 1.9) for the treadmill run (*somewhat hard* to *hard* intensity); 14.2 (SD 1.6) for the embedded run (*somewhat hard* to *hard* intensity), 8.3 (SD 1.6) for the outdoor walk (*extremely light* to *very light* intensity), 13.7 (SD 1.4) for the outdoor run (*somewhat hard* to *hard* intensity), 13.0 (SD 1.7) for using an elliptical machine (*somewhat hard* intensity); 11.3 (SD 1.4) for outdoor biking (*light* to *somewhat hard* intensity); and 14.2 (SD

2.1) for swimming in a lap pool (*somewhat hard* to *hard* intensity).

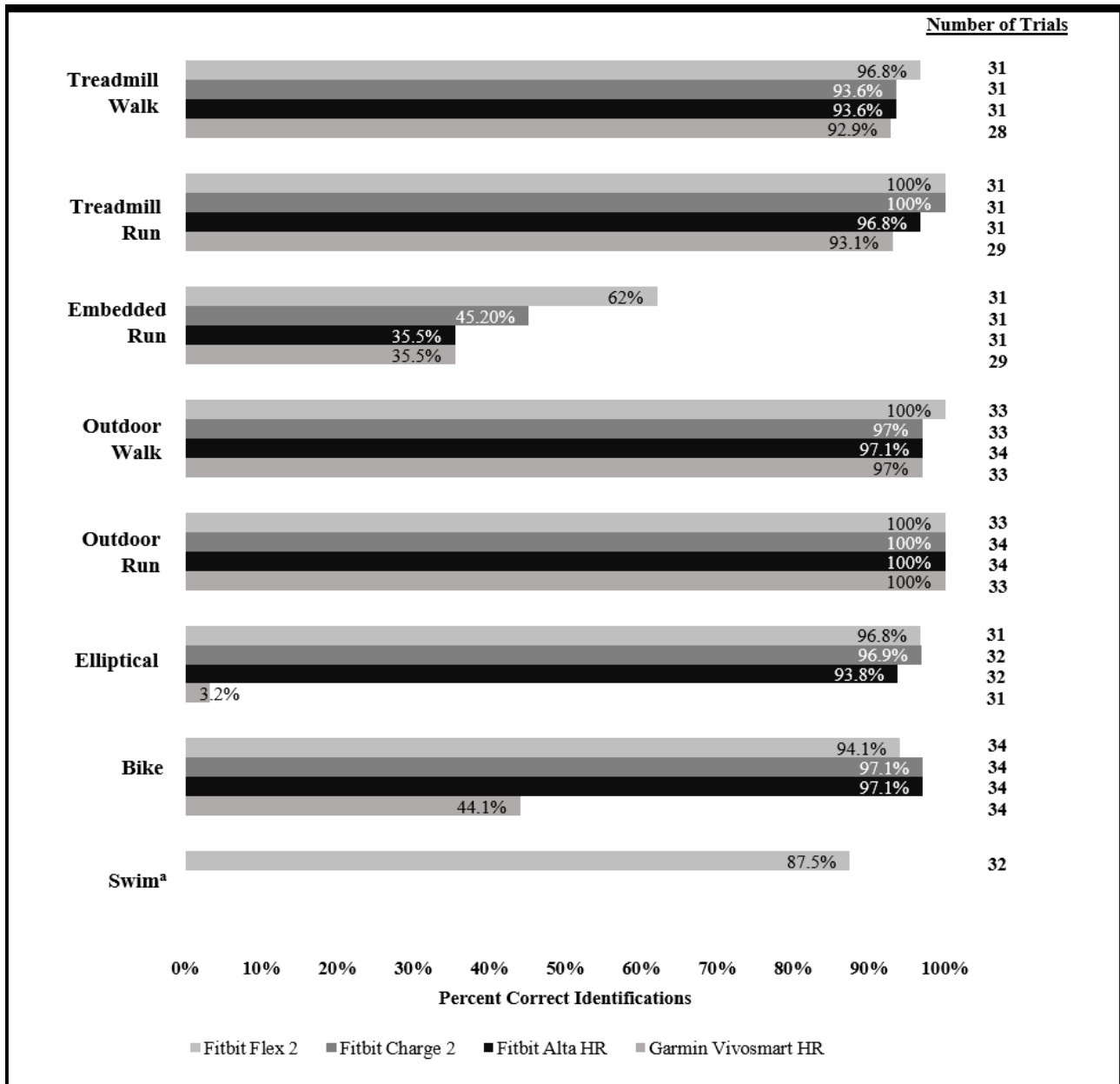
## Activity Identification

Devices were highly successful in recognizing the simple bouts of walking and running on a treadmill or outdoors, struggled to recognize the embedded run, and had more variable success in recognizing the other activities (elliptical trainer, biking outdoors, and swimming; [Figure 1](#)). Overall, the activity with the highest proportion of correct identifications was outdoor running, which was detected for 100% of trials on all devices. Outdoor walking had a rate of correct identifications between 97% and 100%, dependent on tracker brand and model. Although slightly lower, both treadmill walking and running had rates of identification above 92%. Treadmill walking was correctly detected between 93% and 97% of trials dependent on tracker brand and model, and treadmill running was correctly detected between 93% and 100% of trials. The embedded run was detected less often between 36% and 62% of trials. As we were interested in the trackers' ability to detect a run when it was preceded or followed by a short bout of walking, the correct designation for this activity was *run*.

Elliptical use was correctly detected in 91% to 93% of trials for Fitbit devices, but only 3% of trials by the Garmin device. Similarly, biking had a high rate of correct identifications by the Fitbit devices, between 94% and 97% of trials, and a lower recognition rate of 44% of trials by the Garmin device. Swimming was correctly detected for 88% of trials on the Fitbit Flex 2 device. Although we requested the use of front crawl stroke for the duration of the swim module, 5 participants switched to other strokes during the swim. When these participants were excluded from the analysis, the Fitbit Flex 2 had a correct recognition rate of 85% for 27 trials.

The types of observed misclassifications are shown in a confusion matrix ([Table 2](#)). A total of 9 misclassifications occurred for the Garmin device during the elliptical activity, the most misclassifications of any activity. A smaller number of misclassifications were observed for treadmill walking and running, outdoor walking, and swimming. The 2 misclassifications that occurred for the outdoor walking activity were misclassified for the same participant.

Figure 1. Percent correct identifications of eight activities by four wearable trackers.



**Table 2.** Confusion matrix of activity identifications by actual activity type and device.

Activity and device	Identifications		Misclassified as a
	Correct	Erroneous	
<b>Treadmill walk</b>			
Fitbit Flex 2	30	1	Run (1)
Fitbit Charge 2	31	0	— <sup>a</sup>
Fitbit Alta HR	31	0	—
Garmin Vivosmart HR	28	0	—
<b>Treadmill run</b>			
Fitbit Flex 2	31	0	—
Fitbit Charge 2	31	0	—
Fitbit Alta HR	30	1	Aerobic workout (1)
Garmin Vivosmart HR	29	0	—
<b>Embedded run</b>			
Fitbit Flex 2	31	0	—
Fitbit Charge 2	31	0	—
Fitbit Alta HR	31	0	—
Garmin Vivosmart HR	29	0	—
<b>Outdoor walk</b>			
Fitbit Flex 2	33	0	—
Fitbit Charge 2	32	1	Elliptical (1)
Fitbit Alta HR	33	1	Elliptical (1)
Garmin Vivosmart HR	33	0	—
<b>Outdoor run</b>			
Fitbit Flex 2	33	0	—
Fitbit Charge 2	34	0	—
Fitbit Alta HR	34	0	—
Garmin Vivosmart HR	33	0	—
<b>Elliptical</b>			
Fitbit Flex 2	31	0	—
Fitbit Charge 2	32	0	—
Fitbit Alta HR	32	0	—
Garmin Vivosmart HR	22	9	Fitness (9); Other (6); Run (2)
<b>Swim</b>			
Fitbit Flex 2	28	4	Aerobic workout (2); Walk (2)

<sup>a</sup>No activity misclassifications for this device/activity combination.

### Duration Accuracy

Results of the duration accuracy analysis show that detected mean duration was less than 4 min over or under the actual activity duration, except from the treadmill run for all Fitbit trackers, which was overestimated by a mean of 20 min (Figure 2). As the activity bouts were relatively short, even a few minutes of deviation in measurement of activity duration resulted in a substantial MAPE (Table 3). MAPEs were lower for all devices for treadmill walking (7% to 7.9%), outdoor

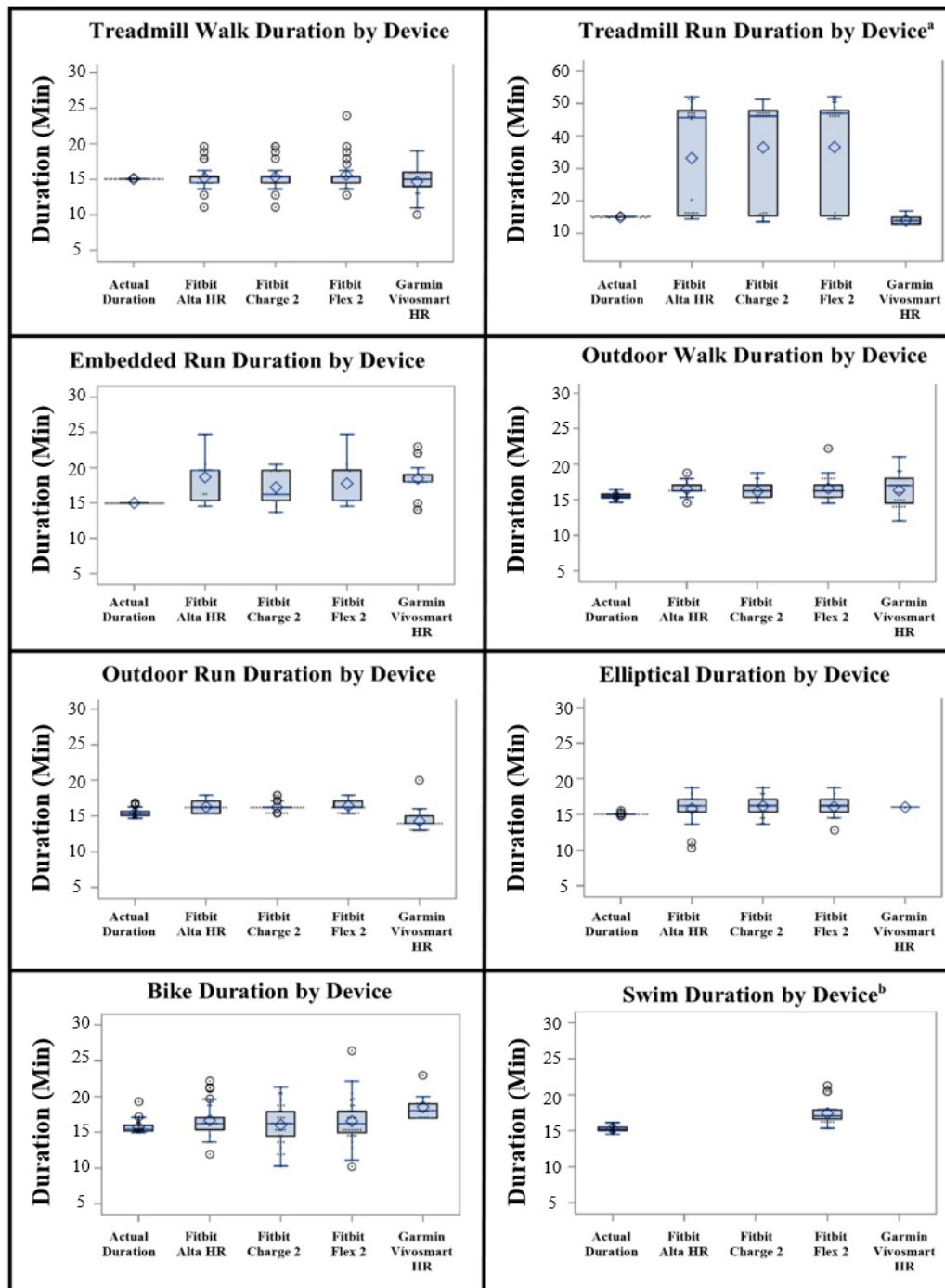
walking (4.9% to 11.8%), and outdoor running (5.6% to 9.6%) than other activities, including treadmill running (8.7% to 144.8%), the embedded treadmill run preceded and followed by a walk (23.6% to 28.9%), an elliptical trainer (9.7% to 13%), biking (9.5% to 17.7%), or swimming (26.9%). Box plots for recorded duration show a spread in time for activities that were completed outdoors or in the pool, as participants needed to return to the activity start site (outdoors) or to the pool edge at the end of the activity. Thus, actual time fluctuated slightly

around 15 min. All devices estimated treadmill walking within 1 min of the actual duration.

The Garmin device underestimated treadmill running duration by approximately 1 min. The embedded run duration was overestimated by a mean of 3 min by all devices. Outdoor walking was overestimated by a mean of 1 min by all devices. Outdoor running was overestimated by a mean of 1 min by Fitbit devices and was underestimated by about 1 min by the

Garmin tracker. Using an elliptical and biking were overestimated by a mean of 1 min by all devices. Swimming was overestimated by approximately 2 min by the Fitbit Flex 2. The Garmin device was the only wearable tracker to underestimate activity duration, and when activity was overestimated by other devices, it was overestimated by less than 4 min on average by all devices, except for the treadmill run, which was overestimated by a mean of 20 min.

**Figure 2.** Box plots of the actual duration of each activity, compared to the duration estimated by the automatic activity recognition feature on each of the four trackers.



**Table 3.** Mean absolute percent error of duration measures by device and activity.

Activity type	Device			
	Fitbit			Garmin Vivosmart HR
	Flex 2	Charge 2	Alta HR	
Treadmill walk	7.9	7	7.6	7.7
Treadmill run	139.4	115.4	144.8	8.7
Embedded run	23.6	24.8	28.9	23.9
Outdoor walk	8.1	6.9	4.9	11.8
Outdoor run	6.6	5.6	5.6	9.6
Elliptical	9.7	10.2	10.2	13
Bike	13.3	9.5	12.9	17.7
Swim	26.9	— <sup>a</sup>	—	—

<sup>a</sup>No data acquired as Only the Fitbit Flex 2 is waterproof and can measure swimming.

## Discussion

### Principal Findings

The most important findings of this study are that the trackers correctly identified most of the common activity types tested in a controlled setting, especially walking and running, and that the trackers had variable success in identifying the duration of the performed activities. This study measured the accuracy of the automatic activity-type recognition and duration feature on 4 wearable activity trackers during walking, running, using an elliptical trainer, biking, and swimming. Although there are countless permutations of possible physical activity behavior (eg, combinations of intensity, duration, setting, and stops and starts), the goal of this study was to conduct a preliminary laboratory-based test of multiple common activity types. All tested trackers had better recognition of ambulatory-based activities such as walking or running as opposed to biking, using an elliptical, or swimming. This finding is consistent with previous validation studies that show wearable activity trackers generally have high correlations of step counts when compared with the criterion during ambulatory activity [23]. Although this study did not aim to validate step counts (which have already been well studied) [23], walking and running were the only activities tested that involved stepping, which may bolster the trackers' detection capability.

Outdoor walking was misclassified as using an elliptical once by both the Fitbit Charge 2 and Fitbit Alta HR. These misclassifications occurred for the same participant and may be because of the participant having an arm movement pattern that mimics an elliptical arm handle pattern during walking. The high misclassification rate of using an elliptical by the Garmin device is concerning and indicates that it is likely unsuitable for apps where participants or patients may engage in use of elliptical machines as a regular part of their exercise regimen. With regard to swimming, although the Fitbit brand recommends swimming freestyle consistently for best detection results, it also claims the Fitbit Flex 2 can measure the backstroke, breaststroke, and butterfly [24]. Although we did not aim to measure strokes other than freestyle, it should be noted that the device did correctly identify all 5 of the swim

trials that experienced a switch in stroke. This is important as stroke switching represents a more natural pattern than swimming a single stroke for an entire bout of activity, suggesting the tracker may be robust for capturing mixed-stroke swimming. However, because of the very small number of trials in which multiple strokes are used, additional data are needed to confirm whether this is true.

The detected mean duration was less than 4 min over or under the actual duration of the activity for outdoor walking and running, the embedded run, using an elliptical, outdoor biking, and swimming. Of these activities, MAPEs were the lowest for outdoor walking and running and treadmill walking. The treadmill run duration for all Fitbit trackers was considerably overestimated by an average of 15 min. The MAPE for treadmill running was also noticeably larger than other activities falling between 115.4% and 144.8%. The considerable overestimation was likely because of the trackers failing to recognize the 10 min of stationary rest after the treadmill run and including some duration of the embedded run (walk, run, walk). The addition of time occurred during 16 trials detected by the Fitbit Alta HR, 20 trials detected by Fitbit Charge 2, and 20 trials detected by the Fitbit Flex 2. If confirmed by future studies, this may indicate that trackers need improvement before they are suitable for satisfactorily identifying and measuring exercise sessions comprising intermittent or combined activities. This is a particular concern as many free-living activities are performed in short bouts and adjacent to other movements; thus, our findings suggest that at least for short bouts of activity (approximately 15 min), current trackers are prone to substantial error in duration estimates.

Similar results were observed for most activities by the Fitbit devices, regarding both activity detection and duration. The 3 Fitbit devices yielded activity detection results that were at most 3.2% different from one another for any ambulatory activity, excluding the embedded run, and had identical results for the outdoor run. This similarity is also observed for the nonambulatory activities. The Fitbit devices yielded results that were, at most, 3.1% different from one another (at least 0.1%) for using an elliptical and, at most, 3% different or identical for outdoor biking. When comparing the Fitbit brand MAPEs, the



treadmill run activity shows the most variability in results with the Fitbit Alta HR and Fitbit Flex 2 differing by 29.4% and with all other activities having identical results or differing by, at most, 5.3%. It is likely that these devices use similar, if not identical, classification algorithms, as they are developed and manufactured by the same company. It is important to consider that future Fitbit devices may continue to yield similar results across devices. If the use of identical technology across devices is confirmed by the company, this may eliminate the need for testing multiple devices in a single study.

This paper provides, to our knowledge, the first data on the automatic activity-type recognition feature in wearable trackers. A major strength is the inclusion of multiple common activities in different contexts (eg, walking outdoors and on a treadmill). The use of a standardized protocol is both a strength and a limitation. As no previous data on this feature are available, we chose a standardized protocol because it provides the ability to directly observe the type and duration of activity participation. However, it is unknown how well the results of this structured experimental study would generalize to free-living settings. Individuals who use wearable activity trackers tend to wear them all day, during which a wide variety of movements and activities are performed. By contrast, in a laboratory-based protocol, the devices are worn only during the study visit and are removed after completion of the final activity in each module. This removal may aid the tracker in determining the stop time of those activities. A second limitation is that we did not attempt to determine how the devices' accuracy might be affected by the duration, speed, or intensity of an activity. Our goal was to test multiple activity types, both indoor and outdoor, to provide a preliminary test of the activity recognition function. Although it is never possible to test all possible combinations of activity type, duration, setting, speed, and

intensity, additional research will provide substantial insight on how these factors generally affect accuracy of the automatic activity recognition feature.

Our results add to the evidence base of previous validation studies that have examined the accuracy of other wearable tracker features, such as steps, which show high correlation between criterion and device [13,14], energy expenditure, which is generally underestimated when compared with the criterion [15], and heart rate, which is more accurate at rest than while exercising [16-18]. Although wearable trackers do not correctly measure all activity metrics perfectly for every trial, their ability to provide objective measures of some aspects of activity is a great advancement from even 10 years ago, with technology likely to improve rapidly with time.

## Conclusions

Wearable activity trackers correctly identified the type of isolated bouts of some common physical activities, although accuracy decreased substantially for certain activities that were adjacent (or nearly so) with one another, and 1 device had very poor detection on the elliptical trainer. Future directions for research on wearable trackers include testing the recognition capabilities of more and complex permutations of activity participation, including aerobic workouts, such as kickboxing or high intensity interval training, sports such as basketball and soccer, slower walking or light-intensity activities (of particular interest for chronic disease populations and older adults), intermittent activities such as jogging interrupted by stoplights, and other free-living activity patterns at a variety of exercise intensities. Owing to the nature of the rapidly evolving wearable technology market, future directions also include testing the automatic activity detection capability in new devices as they are released to the public.

## Acknowledgments

LCB's time is funded by the National Cancer Institute (1K07CA178870). The authors would like to thank the Physical Activity Epidemiology undergraduate students (Bianca Briones and Rachel Schueller) and the study participants for their contributions to this study.

## Conflicts of Interest

None declared.

## References

1. Lee I, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, Lancet Physical Activity Series Working Group. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet* 2012 Jul 21;380(9838):219-229 [FREE Full text] [doi: [10.1016/S0140-6736\(12\)61031-9](https://doi.org/10.1016/S0140-6736(12)61031-9)] [Medline: [22818936](https://pubmed.ncbi.nlm.nih.gov/22818936/)]
2. Arem H, Moore SC, Patel A, Hartge P, Berrington DG, Viswanathan K, et al. Leisure time physical activity and mortality: a detailed pooled analysis of the dose-response relationship. *JAMA Intern Med* 2015 Jun;175(6):959-967 [FREE Full text] [doi: [10.1001/jamainternmed.2015.0533](https://doi.org/10.1001/jamainternmed.2015.0533)] [Medline: [25844730](https://pubmed.ncbi.nlm.nih.gov/25844730/)]
3. Blair SN, Brodney S. Effects of physical inactivity and obesity on morbidity and mortality: current evidence and research issues. *Med Sci Sports Exerc* 1999 Nov;31(11 Suppl):S646-S662. [Medline: [10593541](https://pubmed.ncbi.nlm.nih.gov/10593541/)]
4. Physical Activity Guidelines Advisory Committee. Physical Activity Guidelines Advisory Committee report, 2008. To the Secretary of Health and Human Services. Part A: executive summary. *Nutr Rev* 2009 Feb;67(2):114-120. [doi: [10.1111/j.1753-4887.2008.00136.x](https://doi.org/10.1111/j.1753-4887.2008.00136.x)] [Medline: [19178654](https://pubmed.ncbi.nlm.nih.gov/19178654/)]
5. Piercy KL, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska DA, et al. The physical activity guidelines for Americans. *J Am Med Assoc* 2018 Jan 20;320(19):2020-2028. [doi: [10.1001/jama.2018.14854](https://doi.org/10.1001/jama.2018.14854)] [Medline: [30418471](https://pubmed.ncbi.nlm.nih.gov/30418471/)]

6. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc* 2008 Jan;40(1):181-188. [doi: [10.1249/mss.0b013e31815a51b3](https://doi.org/10.1249/mss.0b013e31815a51b3)] [Medline: [18091006](https://pubmed.ncbi.nlm.nih.gov/18091006/)]
7. Statista. 2019. Number of Fitbit devices sold worldwide from 2010 - 2018 (in 1,000s) URL: <https://www.statista.com/statistics/472591/fitbit-devices-sold/> [accessed 2019-04-13] [WebCite Cache ID 77diTSBYZ]
8. Internicola D. Scientific American. 2014 Dec 22. Activity trackers get smarter at measuring your fitness URL: <https://www.scientificamerican.com/article/activity-trackers-get-smarter-at-measuring-your-fitness/> [accessed 2019-04-14] [WebCite Cache ID 77dilpbOK]
9. Cadmus-Bertram LA, Marcus BH, Patterson RE, Parker BA, Morey BL. Randomized trial of a Fitbit-based physical activity intervention for women. *Am J Prev Med* 2015 Sep;49(3):414-418. [doi: [10.1016/j.amepre.2015.01.020](https://doi.org/10.1016/j.amepre.2015.01.020)] [Medline: [26071863](https://pubmed.ncbi.nlm.nih.gov/26071863/)]
10. Cadmus-Bertram L, Marcus BH, Patterson RE, Parker BA, Morey BL. Use of the Fitbit to measure adherence to a physical activity intervention among overweight or obese, postmenopausal women: self-monitoring trajectory during 16 weeks. *JMIR Mhealth Uhealth* 2015;3(4):e96 [FREE Full text] [doi: [10.2196/mhealth.4229](https://doi.org/10.2196/mhealth.4229)] [Medline: [26586418](https://pubmed.ncbi.nlm.nih.gov/26586418/)]
11. Phillips SM, Cadmus-Bertram L, Rosenberg D, Buman MP, Lynch BM. Wearable technology and physical activity in chronic disease: opportunities and challenges. *Am J Prev Med* 2018 Jan;54(1):144-150. [doi: [10.1016/j.amepre.2017.08.015](https://doi.org/10.1016/j.amepre.2017.08.015)] [Medline: [29122356](https://pubmed.ncbi.nlm.nih.gov/29122356/)]
12. Gualtieri L, Rosenbluth S, Phillips J. Can a free wearable activity tracker change behavior? The impact of trackers on adults in a physician-led wellness group. *JMIR Res Protoc* 2016 Nov 30;5(4):e237 [FREE Full text] [doi: [10.2196/resprot.6534](https://doi.org/10.2196/resprot.6534)] [Medline: [27903490](https://pubmed.ncbi.nlm.nih.gov/27903490/)]
13. Diaz KM, Krupka DJ, Chang MJ, Peacock J, Ma Y, Goldsmith J, et al. Fitbit®: an accurate and reliable device for wireless physical activity tracking. *Int J Cardiol* 2015 Apr 15;185:138-140. [doi: [10.1016/j.ijcard.2015.03.038](https://doi.org/10.1016/j.ijcard.2015.03.038)] [Medline: [25795203](https://pubmed.ncbi.nlm.nih.gov/25795203/)]
14. Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data. *J Am Med Assoc* 2015 Feb 10;313(6):625-626. [doi: [10.1001/jama.2014.17841](https://doi.org/10.1001/jama.2014.17841)] [Medline: [25668268](https://pubmed.ncbi.nlm.nih.gov/25668268/)]
15. O'Driscoll R, Turicchi J, Beaulieu K, Scott S, Matu J, Deighton K, et al. How well do activity monitors estimate energy expenditure? A systematic review and meta-analysis of the validity of current technologies. *Br J Sports Med* 2018 Sep 07 Epub ahead of print(forthcoming). [doi: [10.1136/bjsports-2018-099643](https://doi.org/10.1136/bjsports-2018-099643)] [Medline: [30194221](https://pubmed.ncbi.nlm.nih.gov/30194221/)]
16. Cadmus-Bertram L, Gangnon R, Wirkus EJ, Thraen-Borowski KM, Gorzelitz-Liebhauser J. The accuracy of heart rate monitoring by some wrist-worn activity trackers. *Ann Intern Med* 2017 Apr 18;166(8):610-612. [doi: [10.7326/L16-0353](https://doi.org/10.7326/L16-0353)] [Medline: [28395305](https://pubmed.ncbi.nlm.nih.gov/28395305/)]
17. Wang R, Blackburn G, Desai M, Phelan D, Gillinov L, Houghtaling P, et al. Accuracy of wrist-worn heart rate monitors. *JAMA Cardiol* 2017 Jan 01;2(1):104-106. [doi: [10.1001/jamacardio.2016.3340](https://doi.org/10.1001/jamacardio.2016.3340)] [Medline: [27732703](https://pubmed.ncbi.nlm.nih.gov/27732703/)]
18. Wallen MP, Gomersall SR, Keating SE, Wisløff U, Coombes JS. Accuracy of heart rate watches: implications for weight management. *PLoS One* 2016;11(5):e0154420 [FREE Full text] [doi: [10.1371/journal.pone.0154420](https://doi.org/10.1371/journal.pone.0154420)] [Medline: [27232714](https://pubmed.ncbi.nlm.nih.gov/27232714/)]
19. Fitbit. 2017. Fitbit SmartTrack - Auto Exercise Recognition URL: <https://www.fitbit.com/smarttrack> [accessed 2019-04-14] [WebCite Cache ID 77djOKMdF]
20. Garmin. 2017. Understanding Move IQ URL: [https://support.garmin.com/en-US/?faq=zgFpy8MShkArqAxCug5wC6&searchType=noProduct&utm\\_source=faqSearch](https://support.garmin.com/en-US/?faq=zgFpy8MShkArqAxCug5wC6&searchType=noProduct&utm_source=faqSearch) [accessed 2019-04-14] [WebCite Cache ID 77dkDvIQr]
21. Borg GA. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc* 1982;14(5):377-381. [Medline: [7154893](https://pubmed.ncbi.nlm.nih.gov/7154893/)]
22. Chen M, Fan X, Moe S. Criterion-related validity of the Borg ratings of perceived exertion scale in healthy individuals: a meta-analysis. *J Sports Sci* 2002 Nov;20(11):873-899. [doi: [10.1080/026404102320761787](https://doi.org/10.1080/026404102320761787)] [Medline: [12430990](https://pubmed.ncbi.nlm.nih.gov/12430990/)]
23. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. *Int J Behav Nutr Phys Act* 2015;12(1):159 [FREE Full text] [doi: [10.1186/s12966-015-0314-1](https://doi.org/10.1186/s12966-015-0314-1)] [Medline: [26684758](https://pubmed.ncbi.nlm.nih.gov/26684758/)]
24. Fitbit. 2019 Apr 03. How do I track swims with my Fitbit device? URL: [https://help.fitbit.com/articles/en\\_US/Help\\_article/2087](https://help.fitbit.com/articles/en_US/Help_article/2087) [accessed 2019-04-13] [WebCite Cache ID 77dj2ESRs]

---

## Abbreviations

**MAPE:** mean absolute percent error

**RPE:** rating of perceived exertion

---

*Edited by G Eysenbach; submitted 30.01.19; peer-reviewed by J Sasaki, M Boukhechba, K Winfree, S Gomersall, MS Aslam, C Fincham, R Islam; comments to author 27.02.19; revised version received 13.03.19; accepted 23.03.19; published 23.05.19*

*Please cite as:*

*Dorn D, Gorzelitz J, Gangnon R, Bell D, Koltyn K, Cadmus-Bertram L*

*Automatic Identification of Physical Activity Type and Duration by Wearable Activity Trackers: A Validation Study*

*JMIR Mhealth Uhealth 2019;7(5):e13547*

*URL: <http://mhealth.jmir.org/2019/5/e13547/>*

*doi: [10.2196/13547](https://doi.org/10.2196/13547)*

*PMID: [31124470](https://pubmed.ncbi.nlm.nih.gov/31124470/)*

©Diana Dorn, Jessica Gorzelitz, Ronald Gangnon, David Bell, Kelli Koltyn, Lisa Cadmus-Bertram. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 23.05.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.