



Validity and Wearability of Consumer-based Fitness Trackers in Free-living Children

KEITH BRAZENDALE^{‡1}, LINDSAY DECKER^{†1}, ETHAN T. HUNT^{†1}, MICHELLE W. PERRY^{†1}, ALLISON B. BRAZENDALE^{‡2}, R. GLENN WEAVER^{†1}, and MICHAEL W. BEETS^{†1}

¹Department of Exercise Science, Arnold School of Public Health, University of South Carolina, Columbia, South Carolina, USA; ²Autism Academy of South Carolina, Columbia, South Carolina, USA

†Denotes graduate student author, ‡Denotes professional author

ABSTRACT

International Journal of Exercise Science 12(5): 471-482, 2019. Over the past decade wearable fitness trackers (WFTs) have grown in popularity with more recent versions able to capture the pulse rate noninvasively on the wrist of the wearer. Most of evidence on the validity of WFTs have explored adults in clinical settings. Thus, the purpose of this study is to 1) evaluate the validity of a wrist-placed WFT in determining heart rate, and 2) examine the wear compliance of a wrist-placed WFT, in children in free-living settings. In study 1, 19 children (5-12yrs) wore a Fitbit Charge HR[®] and a Polar chest strap heart rate (HR) monitor for 2 hours while performing sedentary-to-vigorous activities at a holiday camp in December 2016. In study 2, 20 children with mild developmental disabilities (8-13yrs) were asked to wear a Fitbit Alta HR[®] during summer 2017. In study 1, mean absolute percent difference between the WFT HR and criterion was 6.9%. Overall, >75% of WFT HRs were within 5-10% of the criterion. Bland Altman plots indicated a moderate-to-high level of agreement between the WFT and criterion (mean difference 4.1%; Limits of Agreement 26.8, -18.5%). In study 2, participants had the device in their possession for 43 days (SD±14, range 14 - 56 days) and wore it on 67% of those days (range: 20 - 96%) for at least 10 hours/day. Preliminary evidence suggests that WFTs can provide comparable HR estimates to a criterion field-based measure and children can wear WFTs for extended monitoring periods in free-living settings.

KEY WORDS: Validation, child, activity, heart rate, Fitbit, technology

INTRODUCTION

Over the last 10 years, wearable fitness trackers (WFTs) have grown in popularity. More recent versions of WFTs have incorporated photoplethysmography that captures the pulse rate noninvasively on the wrist of the wearer (31). This technology allows for the evaluation of pulse rate as a mark of intensity of physical activity (PA). Recently, several studies in adults

that examined the validity of the pulse rate features of WFTs showed the heart rate (HR) information derived from WFTs is strongly related to the heart rate obtained from scientific-grade heart rate sensors placed around the chest (1, 6). These data indicate the HR feature of WFTs could be used to monitor the intensity of PA of individuals under free-living conditions. The majority of validation studies on WFTs, however, have been performed with college-age young adults, older adults, or clinical populations (2, 15, 19). Studies on the validity of HR from WFT with children are lacking, specifically studies investigating the validity and 'wearability' of these devices in free-living settings (10). Such studies would serve as an important contribution to the science of monitoring youth PA given other advantageous features inherent in many WFTs on the market.

From a measurement standpoint, one of the more attractive features of WFTs is the 'wearability' of the devices which are designed to be worn continuously 24/7 as part of day-to-day life. This feature may improve compliance with wearing a PA monitoring device (35). Lack of compliance with wearing scientific grade motion sensors, such as the ActiGraph accelerometer, for the necessary number of days and number of hours per day is an issue commonly reported in the youth physical activity literature (36). Protocols usually call for wearing a monitor for 7 days, with the hope that 4 of these days (with one being a weekend day) meet the wear time criteria per day (i.e., 6 to 10 hours minimum)(34). A study (38) of over 2,000 9-15 year-olds showed that with a standard 7-day protocol, less than 34% of the participants had complete 7-day data, with a minimum of 6 hours wear/day. This was similar to the wear compliance reported in the NHANES objective assessment of PA via ActiGraph accelerometry (33). These findings suggest that when wear compliance is low, the standard 7-day protocol now needs to be expanded to a minimum of 3 to 4 weeks to reliably measure children's PA.

Should the photoplethysmography feature of WFTs be a valid indicator of HR and children are willing to wear such a device for extended monitoring timeframes, this could greatly increase the field's understanding of youth physical activity behaviors (21, 31). In addition, combining HR with accelerometry has been shown to provide the greatest precision in estimating physical activity and energy expenditure in children (4, 8). Finally, there are a lack of studies examining the utility of WFTs with children in free-living settings (10). Therefore, the purpose of this two-part study was, 1) to examine the validity of WFTs as a measure of heart rate (HR) in children when compared to a criterion measure, and 2) to illustrate the wear compliance from children wearing WFTs over extended timeframes in free-living conditions.

METHODS

Two separate studies were conducted to 1) assess the validity of HR obtained from a WFT (Study 1), and 2) illustrate wear compliance of a wrist-worn WFT in free-living conditions. Study 1 participants were typically-developing elementary school-aged children attending a two-week holiday camp during December, 2016. Study 2 participants were elementary school-aged children with no physical disabilities attending an eight-week summer day camp for

children with developmental disabilities during June and July, 2016. Study 1 and Study 2 procedures were approved by the lead authors institutional review board (IRB).

Participants

Study 1: Children (Mean age: 8.0, ± 1.8 years, 46% female, 60% non-Hispanic white) attending a two-week holiday camp in a southeastern U.S. state were invited to be part of study 1. Inclusion criteria consisted of 1) a child without any physical disability that would prohibit participation in activity and 2) a child with an age between 5-12 years. Parental consent was obtained from the parent/guardian, and verbal assent obtained from the child, prior to the first day of data collection.

Protocol

Study 1: Data collection took place between 7:00am - 12:00pm (Mon-Fri). The morning was split in to two 2-hour data collection segments 7:30 - 9:30am, and 10:00am - 12:00pm. In each 2-hour segment, four children were each given: 1) a Fitbit Charge HR[®] to wear on their non-dominant wrist, 2) a Polar H7[®] watch on their dominant wrist, and, 3) a Polar H7[®] (Polar Electro Inc., Lake Success, NY, USA) telemetry chest strap affixed around the bottom of their sternum. The placement of these devices was conducted by, and took place under, the supervision of a gender-matched trained research assistant and camp counselor, respectively. A second trained research assistant noted the time that the devices were placed on (TIME ON), and removed (TIME OFF), after the two-hour segment. Prior to data collection, the time for the Fitbit Charge HR[®] and the Polar H7[®] watch were calibrated to the nearest second. This protocol was followed for the second morning data collection segment of the same day (10:00am - 12:00pm), but with a different set of four children. For the duration of each 2-hour segment, children played a variety of activities that consisted of staff-led structured games (e.g., tag, basketball) and free-play opportunities. This process was repeated for all days (Mon-Fri).

Statistical Analysis

Study 1: Data from the Fitbit Charge HR[®] was downloaded via a third-party research platform, Fitabase[®] (Small Steps Labs LLC, San Diego, CA, USA) and from the Polar H7[®] via manufacturer software. Data was cleaned for the removal of corrupt files due to criterion measure device malfunction and data from both devices were matched for time (second-by-second) according to time-stamps for each activity session and individual child. Descriptive statistics and Pearson correlations were calculated to determine the magnitude of the relationship between the Fitbit Charge HR[®] and Polar H7[®] estimates. The mean absolute percent difference between the WFT and the criterion measure was examined by the following categories <5% different (from criterion), 5-10% different, and >10% different. Bland Altman plots (95% Limits of Agreement, LOA) were performed to compare the accuracy of the WFT to the criterion measure. All analyses were performed using Stata (V.14.1, College Station, TX).

Participants

Study 2: Twenty children (Mean age: 9, ± 1.6 years, 50% girls, 90% non-Hispanic white) with mild development disabilities attending an eight-week summer day camp (Mid-June to August 2017) were invited to be part of study 2. Inclusion criteria consisted of a child 1) without any physical disability, and 2) between 5-15 years. Parental consent was obtained from the parent/guardian, and verbal assent obtained from the child, prior to the first day of data collection.

Protocol

Study 2: The summer day camp took place at a special education school and catered towards children 5 to 15 years old with social skills deficits. Camp operating hours were from 8:30am – 3:30pm Wednesdays and Fridays. Children participating in the study were given a Fitbit Alta HR[®] to wear on their non-dominant wrist on their first day of camp. Children were instructed to wear it on camp and non-camp days morning and night, handing the device back on their last day of camp. The time and date the device was placed on, and removed from, the child's non-dominant wrist was recorded by research staff. During the first day of camp, a trained research assistant gave each parent/guardian of a participating child a unique login and password for their pre-configured Fitbit[®] accounts, instructions for the parent/guardian on how to download the Fitbit[®] app, sync, and charge the device. Each account was linked to an online account (12). Through Fitabase[®], research assistants monitored battery life and frequency of syncing events and sent text reminders to the parent/guardian to address any long periods of non-compliance defined as ≥ 4 days of no syncing/charging.

Statistical Analysis

Study 2: Currently, there are no defined procedures to analyze or interpret Fitbit[®] HR coverage data (3, 10). Each child's individual HR files were downloaded from Fitabase[®] in 1-minute epochs and cleaned to remove the first and last days of data as these were not complete days (i.e. 24 hours). All days the child had the device in their possession was considered 'Wearable Days'. A valid wear-day was defined as any day with at least 10 hours of 60 second epoch HR recordings present. Non-wear time was classified as any 15-minute period on a wear-day where consecutive HR values were missing. Secondary wear time analysis explored day and night wear and non-wear defined as HR recordings occurring between 8:00am – 8:00pm and 8:01pm – 7:59am, respectively. Valid wear-day data was interpreted as the percent of wear-days ranging from 10 to 24 hours/day of 60 second epoch HR recordings present, hereon referred to as 'HR coverage'.

RESULTS

A total of nineteen children provided observations (2-hour segments) for a total of 38,778 1-minute HR data points for comparison in Study 1. Table 1 shows the mean HR beats per minute (bpm) for the WFT and the criterion measure. Overall, the correlation between Fitbit Charge HR[®] and Polar H7[®] was $r = 0.84$, represented by an absolute HR difference of 8.9 bpm, and an absolute percent difference of 6.9% (Table 1.) Figure 1 shows the results of a

Bland-Altman plot where 75% of the HR estimates from the Fitbit device were within $\pm 10\%$ of the criterion measure, with 51% falling within $<5\%$. Bland Altman plots indicated a moderate-to-high level of agreement between the WFT and criterion (mean difference 4.1%; LOA 26.8, -18.5%).

Table 1. Study 1: Heart rate estimates from criterion and consumer wearable fitness tracker for the overall sample and by level of agreement.

	Overall				Absolute Percent Difference											
					0 to $\leq 5\%$				$>5\%$ to $\leq 10\%$				$>10\%$			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Data Coverage (percentage)	100				54.4				22.7				22.9			
Heart Rate (bpm)																
Criterion (Polar Chest Strap)	120.4	21.1	66.0	211.0	118.6	19.4	71.0	199.0	119.6	20.2	76.0	194.0	125.6	24.7	66.0	211.0
FitBit Charge HR	116.2	19.5	71.0	195.0	117.9	19.3	71.0	195.0	115.5	19.2	71.0	187.0	113.0	20.0	71.0	173.0
Heart Rate Difference (bpm)	4.2	11.3	-55.0	97.0	0.7	3.4	-9.0	10.0	4.1	8.6	-16.0	19.0	12.6	18.9	-55.0	97.0
Absolute Difference (bpm)	8.5	8.6	0.0	97.0	2.9	2.0	0.0	10.0	9.2	2.4	5.0	19.0	21.1	8.5	8.0	97.0
Heart Rate Percent Difference (percent)	2.9	9.0	-65.5	46.0	0.5	2.9	-5.5	5.5	3.1	7.1	-10.5	10.5	8.5	15.4	-65.5	46.0
Absolute Difference (percent)	6.9	6.5	0.0	65.5	2.4	1.6	0.0	5.0	7.7	1.4	5.0	10.0	16.7	5.6	10.0	65.5

Note: Heart rate difference calculated as criterion minus FitBit. Heart rate percent difference calculated as heart rate difference divided by criterion.

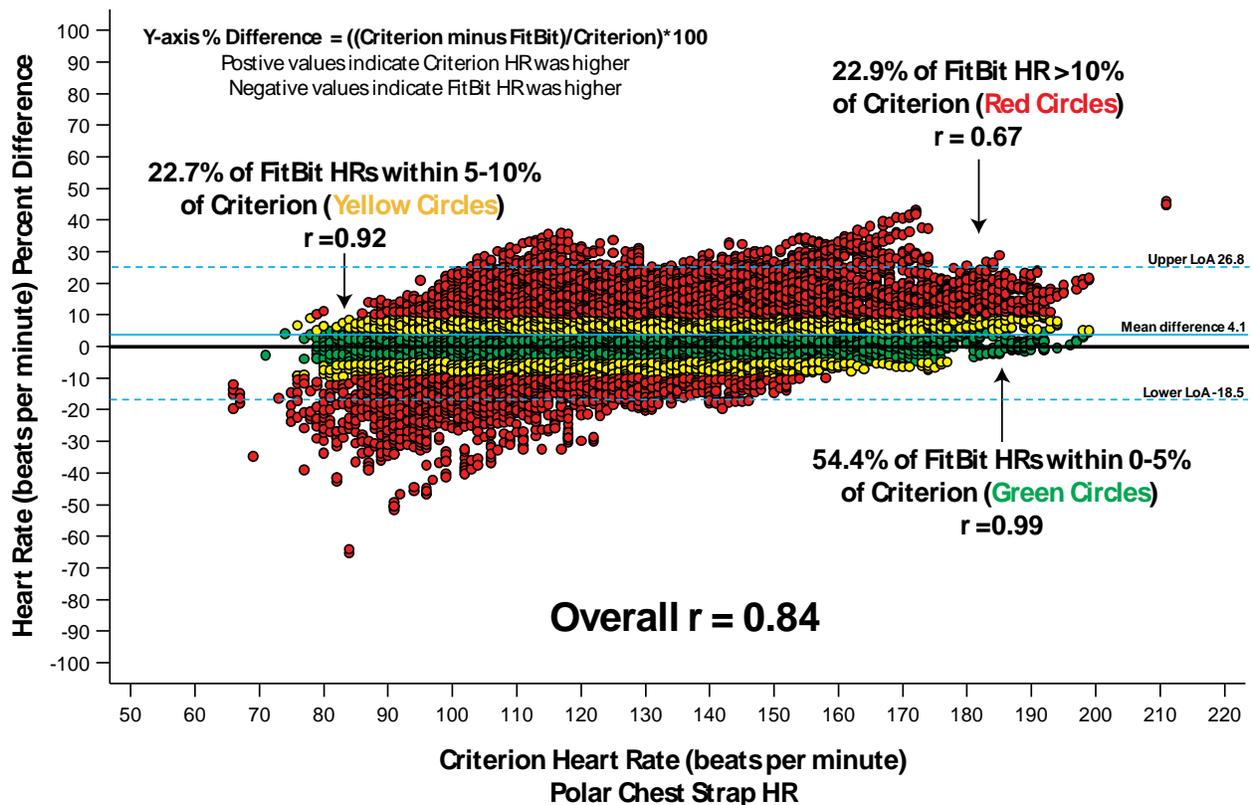


Figure 1. Study 1: Bland-Altman plot of criterion heart rate versus percent difference of Fitbit heart rate.

Table 2. Study 2 Participant Wear time, Non-wear time and Heart rate coverage

Wear time				
	Median	SD (±)	Range	
			Min.	Max.
Child Observation Days				
Wearable Days ^a (n=781)	43	14	14	56
Valid Wear Days ^b (n=477)	24	13	3	52
% of Wearable Days Worn	67	23	20	96
Minutes/Day				
All day (12:00am - 11:59pm)	1,212	431	1	1,440
Day (8:00am - 8:00pm)	703	220	0	720
Night (8:01pm - 7:59am)	601	271	0	720
Non-Wear time				
	Median	SD (±)	Range	
			Min.	Max.
Child Observation Days				
Non-wear days ^c (n=427)	32	10	8	41
Minutes/Day				
All day (12:00am - 11:59pm)	420	422	15	1,440
Day (8:00am - 8:00pm)	195	239	15	705
Night (8:01pm - 7:59am)	390	234	15	735
Heart Rate Coverage				
Coverage Criteria (hours per day)	Percent of Valid Wear Days^d			
	Median	Range		
12	85%	39 - 100%		
14	83%	31 - 100%		
16	77%	13 - 100%		
18	69%	9 - 100%		
20	64%	4 - 96%		
22	59%	4 - 92%		
24	40%	4 - 68%		

^a All possible days Fitbit© was in child’s possession. ^b Days with at least 10 hours of 60-second epoch heartrate recordings.

^c Non-wear time identified as any day with one 15-minute period without a heartrate recording. ^d Values from a fitted regression of individual participants’ (n=20) HR coverage.

The children invited to be part of study 2 (N=20) had the device in their possession for 781 child observation days, representing a median of 43 wearable days (\pm SD 14, Range: 14 - 56 days). Median valid wear days (i.e., \geq 10 hours recording/day) in the sample was 24 (\pm SD 13, Range: 3 - 52 days), representing 67% of the days (SD \pm 23, range 20 - 96% of days) the device was in their possession. Table 2 presents wear time, non-wear time, and HR coverage information for the sample. Median wear time minutes per day (min/day) were 1,212 (\pm SD 431, Range: 1 - 1,440 minute), with higher median wear-time minutes recorded during the day (703 min/day) compared to night (601 min/night). Non-wear time occurred on 427 child observations days (median 32, \pm SD 10, range 8 - 41 days). Median non-wear time was greater at night compared to during the day (390 vs. 195 min/day) (Table 2). The median percent of valid wear days with complete HR coverage from 12 hours/day (85%, range:39 - 100%) to 24 hours/day (40%, range: 4 - 68%) is presented in Table 2.

DISCUSSION

To our knowledge, this is one of the first studies to examine 1) the validity of a WFT as a measure of HR in children compared to a criterion measure and 2) the compliance and coverage from children wearing WFTs over extended timeframes. Results from this preliminary two-part study suggest that WFTs can provide comparable HR estimates to a criterion field-based measure. In addition, a high degree of user compliance and substantial HR coverage was observed in a separate sample of children.

Our findings indicate that WFTs utilizing photoplethysmography, such as Fitbit® devices, are a satisfactory method to monitor activity in children. These findings are encouraging given the advantages WFTs offer. First, HR technology affords researchers the ability to capture a marker of physical activity which is to each child's level of fitness (i.e., resting HR) and can be individualized for intensity classification purposes (e.g., % HR reserve) (27). Second, WFTs have the potential to be highly informative from a wear-time/compliance monitoring perspective (i.e., HR detected vs. no HR detected). Third, WFTs offer online databases synced to the participants device giving researchers the ability to monitor activity remotely thereby reducing the issue of losing large amounts of data when a device that stores data locally (e.g., most scientific-grade accelerometers) is misplaced or malfunctions. Lastly, WFTs are aesthetically designed to be worn and stay worn, and, thus, may lead to a larger number of wear days, and, thus a longer assessment of activity-related markers (10).

There is a lack of literature investigating the validity of HR as captured by a WFT in comparison to a criterion measure in children (10). Previous youth intervention and non-intervention field-based studies have used WFTs, yet have chosen to focus on other health-related features offered by the device, such as step count or total sleep time (16-18, 22). Results from study 1 demonstrate moderate-to-high levels of HR agreement between the Fitbit Charge HR® and the criterion measure. Over 75% of the ~38,000 HR observations were within a 10% difference. Several validation studies using Fitbit® devices and a HR chest strap as a criterion measure have been conducted in samples of young adult (18-38 years old). The majority of

these studies were lab-based (10). These studies reported mean absolute percent error between 3-16% concluding WFTs can produce valid estimates (5, 6, 30, 37). The results herein, albeit from a field-based setting in a younger population, show mean absolute percent error within a similar range (2.4-16.7%), and in accordance with other studies (5, 10), a systematic underestimation of HR as intensity increases from the Fitbit® devices (see Table 1/Figure 1). Given this, researchers must be cautious of misclassification when interpreting HR data using Fitbit® devices, particularly during protocols that have participants spending extended periods of time at high-intensities. Further, future research should explore possible reasons as to why this underestimation of HR with increasing intensity may be occurring. Currently, only a handful of studies have explored potential causes of discrepancies of WFTs and photoplethysmography citing issues such as skin pigment, the position of the wrist, and the design of the wrist strap for appropriate snugness of fit on children's wrists (29). Nonetheless, collectively, our findings indicate that the Fitbit charge HR® produces similar estimates of HR in children in comparison to a criterion measure. Incorporating objective measures that utilize photoplethysmography may prove advantageous when concerned with measuring children's health-enhancing physical activity – such as moderate-to-vigorous PA – as it allows researchers to obtain individually standardized moderate-to-vigorous PA estimates based on each individual child's resting heart rate, a marker of fitness (7, 27).

Results from Study 2 illustrate a high degree of wear compliance from participants. Children wore the device for more than 10 hours per day on 67% of the days they had it in their possession, which is comparable to a recent intervention that used Fitbit® devices in a sample of school-aged children where participants wore the device on the wrist for 8 hours per day or longer for 80% of possible days (9). Median valid wear days in the sample of children in Study 2 was 24 days, which is longer than any previously published field-based study using a wrist-worn WFT with children (<10 days) (10). One study deployed a waist-worn WFT for 124 days in a sample of urban youth (N=24) reporting low levels of adherence, with 19 average wear days (range 1-74 days) per participant (26). Large studies, such as the National Health and Nutrition Examination Survey (NHANES 2011-2012), switched to a wrist-worn accelerometer placement as opposed to the hip-placement in previous studies (NHANES 2003-2006) to increase wear compliance recognizing that longer monitoring periods are associated with more reliable measures of activity (20), and high compliance gives more confidence that the data are representative of daily physical activity (25). The 2011-2012 NHANES reported 70%-80% of participants achieved a median wear time of >20 hours per day for >6 days; substantially higher than the 40%-70% of participants achieving 10 hours per day of wear for >6 days (NHANES 2003-2006) when participants wore the accelerometer on the hip (13). Greater wear compliance demonstrated by wrist-worn WFTs can potentially alleviate concerns researchers may have on how to detect and classify periods of non-wear time, obstacles one must consider when selecting other objective physical activity measurement devices, such as wrist-worn accelerometers (24). For example, with WFTs, photoplethysmography will detect if the individual was compliant and wearing the WFT – as indicated by the presence of a HR – or not. WFTs offer the best of both worlds in the sense that they provide a clear marker of wear (i.e., presence of HR), and they are designed to be worn on the wrist, a more compliant wear

location in children (11). Of note, numerous studies in children and adults have shown that the combination of HR and accelerometry data provides the greatest precision in estimating physical activity and energy expenditure compared to HR or accelerometry alone (4, 8, 28). The number of single devices available to researchers that offer both capabilities of capturing HR and accelerometry data – specifically, access to the raw data (i.e., accelerometer counts) – is limited. Nonetheless, this is important for researchers to consider when selecting a wrist-worn device that meets the needs of their primary outcome.

In addition to wear compliance, the median daily wear-time compliance of children in Study 2, referred to herein as HR coverage, was ~1,200 minutes per day from 691 child observation days (Table 2). Currently, there are no evidence-based criteria on what is acceptable in terms of a valid wear-day for WFTs, specifically for HR coverage (23), and protocols that capture 24-hour activity (physical activity, sedentary time, and sleep), such as in our Study 2. However, the percent of wear days with at least 10 hours of waking wear-time (identified as ≥ 600 HR recordings between 8:00am and 8:00pm) – a common marker of a valid day in field-based assessments of physical activity using accelerometers (36) – was 69%. This is similar to a previous study of adolescents who wore a Fitbit® device for at least 10 hours per day on (33 out of 49 intervention days; 67.3%) (14). These data are greater than existing valid day criteria for objective measures of youth physical activity where, typically, researchers accept 4 valid days from a 7 or 10 day data collection protocol (57% of days or 40% of days, respectively) (34). In addition, Study 2 findings show that setting a more stringent wear criteria of >20 hours per day (≥ 1200 HR recordings per day) still classifies 64% of observation days as valid (Table 2). This illustrates a high degree of wear-time compliance in this sample, and with the recent release of the Canadian 24-hour movement guidelines, offers researchers a timely and attractive alternative to capture 24-hour activity for extended periods of time (i.e., day and night data for multiple weeks, months) (32). Results from our secondary analysis showed that non-wear time occurred more during the night compared to during the day. This was confirmed by conversations with parent/guardians who stated that this was the preferred time to charge the device. Collectively, the ability to conduct a long-term data collection protocol is demonstrated by the adherence of Study 2 parents/guardians to the protocols of downloading the Fitbit® app and syncing and/or charging the device every 3-4 days. Across 20 participants, only 6 parents/guardians were issued a text reminder over summer equating to 16 text messages being sent by study staff (out of a possible 180 text messages).

There are several strengths of this study. First, this is one of the first field-based studies to validate HR obtained from a WFT in a sample of children. Second, this is one of the first field-based studies to explore HR data coverage obtained from a WFT in a sample of children for extended periods of time (>2 weeks). Third, both studies were conducted in field-based settings which captures the free-living sporadic nature of children's physical activity. Lastly, Study 1 used a HR chest strap as the criterion measure, a widely accepted criterion measure for HR. Both studies are not without limitation. Results are only representative of the Fitbit Charge HR® and Alta HR® devices, it is not clear whether similar findings would be obtained from other consumer devices. Although sample size is comparable to other lab and field-

based studies evaluating WFT in children and adults (5, 6, 37), the majority of children in both samples were non-Hispanic white, thus, results may not be generalizable to children from other racial/ethnic groups. Also, children in Study 2 had mild developmental disabilities, however, the authors believe HR coverage and compliance would not decrease had a typically developing sample of children been chosen.

The preliminary evidence presented in this two-part validity study demonstrates that Fitbit® devices 1) provide valid HR estimates in children, 2) provides adequate daily wear-time coverage, and 3) can be used to capture 24-hour data for an extended period of time (>2 weeks) in children. The implications of this are far-reaching. Researchers and practitioners concerned with objectively assessing children's physical activity have a means to do this for longer periods which could greatly increase the field's understanding of youth physical activity behaviors and establish more robust and accurate assessments.

REFERENCES

1. Bai Y, Hibbing P, Mantis C, Welk GJ. Comparative evaluation of heart rate-based monitors: Apple Watch vs Fitbit Charge HR. *J Sports Sci* 36(15): 1-8, 2017.
2. Boudreaux BD, Hebert EP, Hollander DB, Williams BM, Cormier CL, Naquin MR, et al. Validity of wearable activity monitors during cycling and resistance exercise. *Med Sci Sports Exerc* 50(3):624-633, 2018.
3. Bunn JA, Navalta JW, Fountaine CJ, Reece JD. Current state of commercial wearable technology in physical activity monitoring 2015–2017. *Int J Exerc Sci*. 11(7):503-515, 2018.
4. Corder K, Brage S, Wareham NJ, Ekelund U. Comparison of PAEE from combined and separate heart rate and movement models in children. *Med Sci Sports Exerc* 37(10):1761-1767, 2005.
5. Dondzila CJ, Lewis C, Lopez JR, Parker T. Congruent accuracy of wrist-worn activity trackers during controlled and free-living conditions. *Int J Exerc Sci* 11(7):575-584, 2018.
6. Dooley EE, Golaszewski NM, Bartholomew JB. Estimating accuracy at exercise intensities: A comparative study of self-monitoring heart rate and physical activity wearable devices. *JMIR mHealth uHealth* 5(3), 2017.
7. Epstein LH, Paluch RA, Kalakanis LE, Goldfield GS, Cerny FJ, Roemmich JN. How much activity do youth get? A quantitative review of heart-rate measured activity. *Pediatrics* 108(3):e44-e44, 2001.
8. Eston RG, Rowlands AV, Ingledew DK. Validity of heart rate, pedometry, and accelerometry for predicting the energy cost of children's activities. *J Appl Phys* 84(1):362-371, 1998.
9. Evans EW, Abrantes AM, Chen E, Jelalian E. Using novel technology within a school-based setting to increase physical activity: A pilot study in school-age children from a low-income, urban community. *Biomed Res Int* 2017(4): 1-7, 2017.
10. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. *Int J Behav Nutr Phys Act* 12(1):159, 2015.
11. Fairclough S, Noonan R, Rowlands A, Van Hees V, Knowles Z, Boddy L. Wear compliance and activity in children wearing wrist and hip mounted accelerometers. *Med Sci Sports Exerc* 48(2):245-253, 2016.

12. Franzen-Castle L, Dunker T, Chai W, Krehbiel M. Fitbit and fitabase technology: Tracking and evaluating youth physical activity. *J Extension* 55(2): 2TOT5, 2017.
13. Freedson PS, John D. Comment on" estimating activity and sedentary behavior from an accelerometer on the hip and wrist". *Med Sci Sports Exerc* 45(5):962-963, 2013.
14. Gaudet J, Gallant F, Bélanger M. A Bit of Fit: Minimalist intervention in adolescents based on a physical activity tracker. *JMIR mHealth uHealth* 5(7), 2017.
15. Gillinov S, Etiwy M, Wang R, Blackburn G, Phelan D, Gillinov AM, et al. Variable accuracy of wearable heart rate monitors during aerobic exercise. *Med Sci Sports Exerc* 49(8):1697-1703, 2017.
16. Hamari L, Kullberg T, Ruohonen J, Heinonen OJ, Díaz-Rodríguez N, Lilius J, et al. Physical activity among children: objective measurements using Fitbit One® and ActiGraph. *BMC Res Notes* 10(1):161, 2017.
17. Hayes LB, Van Camp CM. Increasing physical activity of children during school recess. *J Appl Behav Anal* 48(3):690-695, 2015.
18. Hooke MC, Gilchrist L, Tanner L, Hart N, Withycombe JS. Use of a fitness tracker to promote physical activity in children with acute lymphoblastic leukemia. *Pediatr Blood Cancer* 63(4):684-689, 2016.
19. Horton JF, Stergiou P, Fung TS, Katz L. Comparison of Polar M600 Optical heart rate and ECG heart rate during exercise. *Med Sci Sports Exerc* 49(12):2600-2607, 2017.
20. Levin S, Jacobs DR, Ainsworth BE, Richardson MT, Leon AS. Intra-individual variation and estimates of usual physical activity. *Ann Epidemiol* 9(8):481-488, 1999.
21. Lewis BA, Napolitano MA, Buman MP, Williams DM, Nigg CR. Future directions in physical activity intervention research: expanding our focus to sedentary behaviors, technology, and dissemination. *J Behav Med* 40(1):112-126, 2017.
22. Meltzer LJ, Hiruma LS, Avis K, Montgomery-Downs H, Valentin J. Comparison of a commercial accelerometer with polysomnography and actigraphy in children and adolescents. *Sleep* 38(8):1323-1330, 2015.
23. Montgomery-Downs HE, Insana SP, Bond JA. Movement toward a novel activity monitoring device. *Sleep Breath* 16(3):913-917, 2012.
24. Oliver M, Badland HM, Schofield GM, Shepherd J. Identification of accelerometer nonwear time and sedentary behavior. *Res Q Exerc Sport* 82(4):779-783, 2011.
25. Rowlands AV, Rennie K, Kozarski R, Stanley RM, Eston RG, Parfitt GC, et al. Children's physical activity assessed with wrist-and hip-worn accelerometers. *Med Sci Sports Exerc* 46(12):2308-2316, 2014.
26. Schaefer SE, Ching CC, Breen H, German JB. Wearing, thinking, and moving: testing the feasibility of fitness tracking with urban youth. *Am J Health Educ* 47(1):8-16, 2016.
27. Scott SN, Thompson DL, Coe DP. The ability of the PACER to elicit peak exercise response in youth [corrected]. *Med Sci Sports Exerc* 45(6):1139-1143, 2013.

28. Silva AM, Santos DA, Matias CN, Judice PB, Magalhães J, Ekelund U, et al. Accuracy of a combined heart rate and motion sensor for assessing energy expenditure in free-living adults during a double-blind crossover caffeine trial using doubly labeled water as the reference method. *Eur J Clin Nutr* 69(1):20, 2015.
29. Spierer DK, Rosen Z, Litman LL, Fujii K. Validation of photoplethysmography as a method to detect heart rate during rest and exercise. *J Med Eng Tech* 39(5):264-271, 2015.
30. Stahl SE, An H-S, Dinkel DM, Noble JM, Lee J-M. How accurate are the wrist-based heart rate monitors during walking and running activities? Are they accurate enough? *BMJ Open Sport Exerc Med* 2(1):e000106, 2016.
31. Tamura T, Maeda Y, Sekine M, Yoshida M. Wearable photoplethysmographic sensors – past and present. *Electronics* 3(2):282-302, 2014.
32. Tremblay MS, Carson V, Chaput J-P, Connor Gorber S, Dinh T, Duggan M, et al. Canadian 24-hour movement guidelines for children and youth: an integration of physical activity, sedentary behaviour, and sleep. *Appl Physiol Nutr Metab* 41(6):S311-S327, 2016.
33. 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* 40(1):181, 2008.
34. Trost SG, Pate RR, Freedson PS, Sallis JF, Taylor WC. Using objective physical activity measures with youth: how many days of monitoring are needed? *Med Sci Sports Exerc* 32(2):426, 2000.
35. Tudor-Locke C, Barreira TV, Schuna JM, Mire EF, Chaput J-P, Fogelholm M, et al. Improving wear time compliance with a 24-hour waist-worn accelerometer protocol in the International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE). *Int J Behav Nutr Phys Act* 12(1):11, 2015.
36. Tudor-Locke C, Johnson WD, Katzmarzyk PT. US population profile of time-stamped accelerometer outputs: impact of wear time. *J Phys Act Health* 8(5):693-698, 2011.
37. Wallen MP, Gomersall SR, Keating SE, Wisløff U, Coombes JS. Accuracy of heart rate watches: implications for weight management. *PloS one* 11(5):e0154420, 2016.
38. Wickel EE. Reporting the reliability of accelerometer data with and without missing values. *PloS one* 9(12):e114402, 2014.

