

2018

# Comparison of Four Fitbit and Jawbone Activity Monitors with a Research-Grade ActiGraph Accelerometer for Estimating Physical Activity and Energy Expenditure

Mary T. Imboden  
*George Fox University*

Michael B. Nelson  
*Ball State University*

Leonard A. Kaminsky  
*Ball State University*

Alexander HK Montoye  
*Ball State University*

Follow this and additional works at: [https://digitalcommons.georgefox.edu/hhp\\_fac](https://digitalcommons.georgefox.edu/hhp_fac)

Part of the [Physical Therapy Commons](#), and the [Sports Sciences Commons](#)

---

## Recommended Citation

Imboden, Mary T.; Nelson, Michael B.; Kaminsky, Leonard A.; and Montoye, Alexander HK, "Comparison of Four Fitbit and Jawbone Activity Monitors with a Research-Grade ActiGraph Accelerometer for Estimating Physical Activity and Energy Expenditure" (2018). *Faculty Publications - Department of Health and Human Performance*. 15.  
[https://digitalcommons.georgefox.edu/hhp\\_fac/15](https://digitalcommons.georgefox.edu/hhp_fac/15)

This Article is brought to you for free and open access by the Department of Health and Human Performance at Digital Commons @ George Fox University. It has been accepted for inclusion in Faculty Publications - Department of Health and Human Performance by an authorized administrator of Digital Commons @ George Fox University. For more information, please contact [arolfe@georgefox.edu](mailto:arolfe@georgefox.edu).

# Comparison of four Fitbit and Jawbone activity monitors with a research-grade ActiGraph accelerometer for estimating physical activity and energy expenditure

Mary T Imboden,<sup>1</sup> Michael B Nelson,<sup>1</sup> Leonard A Kaminsky,<sup>2</sup> Alexander HK Montoye<sup>1</sup>

<sup>1</sup>Clinical Exercise Physiology Program, Ball State University, Muncie, Indiana, USA  
<sup>2</sup>Fisher Institute of Health and Well-Being, Ball State University, Muncie, Indiana, USA

**Correspondence to**  
Alexander HK Montoye,  
Integrative Physiology and Health Science, Department, Alma College, 614 W. Superior, Alma, MI 48801, USA;  
montoyeah@alma.edu

## ABSTRACT

**Background/aim** Consumer-based physical activity (PA) monitors have become popular tools to track PA behaviours. Currently, little is known about the validity of the measurements provided by consumer monitors. We aimed to compare measures of steps, energy expenditure (EE) and active minutes of four consumer monitors with one research-grade accelerometer within a semistructured protocol.

**Methods** Thirty men and women (18–80 years old) wore Fitbit One (worn at the waist), Fitbit Zip (waist), Fitbit Flex (wrist), Jawbone UP24 (wrist) and one waist-worn research-grade accelerometer (ActiGraph) while participating in an 80 min protocol. A validated EE prediction equation and active minute cut-points were applied to ActiGraph data. Criterion measures were assessed using direct observation (step count) and portable metabolic analyser (EE, active minutes). A repeated measures analysis of variance (ANOVA) was used to compare differences between consumer monitors, ActiGraph, and criterion measures. Similarly, a repeated measures ANOVA was applied to a subgroup of subjects who didn't cycle.

**Results** Participants took 3321±571 steps, had 28±6 active min and expended 294±56 kcal based on criterion measures. Comparatively, all monitors underestimated steps and EE by 13%–32% ( $p<0.01$ ); additionally the Fitbit Flex, UP24, and ActiGraph underestimated active minutes by 35%–65% ( $p<0.05$ ). Underestimations of PA and EE variables were found to be similar in the subgroup analysis.

**Conclusion** Consumer monitors had similar accuracy for PA assessment as the ActiGraph, which suggests that consumer monitors may serve to track personal PA behaviours and EE. However, due to discrepancies among monitors, individuals should be cautious when comparing relative and absolute differences in PA values obtained using different monitors.

## INTRODUCTION

Physical activity (PA) reduces risk of obesity, diabetes, hypertension, cardiovascular disease and all-cause mortality.<sup>1</sup> National PA guidelines recommend that adults participate in  $\geq 150$  min/week of moderate-intensity or vigorous-intensity PA as a minimal amount of PA needed to increase or maintain health.<sup>1</sup> Accurate assessment of individuals' PA, including volume, intensity and time, is important.<sup>2</sup>

Accelerometers provide an objective assessment of PA by measuring accelerations of the body and

translating these accelerations into PA variables such as steps, energy expenditure (EE) and time spent in moderate-intensity or vigorous-intensity PA (active minutes). However, accelerometers have mainly been used in research settings and are rarely used by consumers for personal PA tracking.

More recently, consumer-based PA monitors, which use accelerometer technology, have become popular as personal PA tracking tools. In 2013, consumer monitor sales were estimated at \$330 million worldwide, with approximately 87% coming from Fitbit (Fitbit, San Francisco, California, USA) and Jawbone (Jawbone, San Francisco, California, USA) brands.<sup>3</sup> Like research-grade accelerometers, consumer monitors estimate PA variables including steps taken, EE and active minutes. However, consumer monitors offer numerous advantages over most research-grade accelerometers, including real-time feedback, easy synchronisation to smartphone or computer applications, and goal-tracking.

Despite the recent widespread adoption of consumer monitors, little research has compared their accuracy with research-grade accelerometers. Lee *et al* compared the accuracy of consumer monitors (Fitbit One (FO), Fitbit Zip (FZ), Jawbone UP24 (JU)) and one popular research-grade accelerometer, the ActiGraph (AG; ActiGraph, Pensacola, Florida, USA), with indirect calorimetry for estimating EE in a semistructured setting. These researchers found similar accuracy of the consumer monitors to the research-grade accelerometer, with mean absolute per cent error (MAPE) of 10%–12% for the consumer monitors and 12.6% for the research-grade accelerometer compared with indirect calorimetry.<sup>4</sup> In a similar investigation, Bai *et al* assessed the accuracy of the Fitbit Flex (FF), JU, and AG against indirect calorimetry for measuring EE in a semistructured setting, finding that the FF, JU, and AG all had MAPE of  $<20\%$  for EE measurements.<sup>5</sup> Finally, Murakami *et al* assessed the accuracy of 12 activity monitors, including FF, JU, and AG for measuring EE in a free-living setting compared with doubly labelled water. Estimates from the 12 devices ranged from 590 to 69 kcal/day lower than the doubly labelled water measure.<sup>6</sup> While these studies provide insight into EE prediction, they did not assess accuracy of steps or active minute estimates. Storm *et al* compared seven activity monitors, including two research-grade accelerometers, the MoveMonitor and activPAL (PAL Technologies,

Glasgow, UK). All monitors underestimated steps compared with the criterion measure; however, the MoveMonitor accelerometer had the best performance with less than 2% error at all walking speeds. Two recent studies showed moderate or strong correlations between consumer monitors and the AG for step counting ( $r=0.80-0.91$ ), EE ( $r=0.74-0.81$ ) and active minutes ( $r=0.52-0.91$ ) under free-living conditions. However, these studies lacked a criterion measure, so accuracy of these devices could not be determined.<sup>7,8</sup>

Due to the widespread use of consumer monitors and other PA and health monitoring tools, it is important to gain understanding on how consumer monitors compare with a popularly used research-grade accelerometer in terms of accuracy of PA measurement and to make comparisons of studies using different types and brands of monitors. Thus, we compared four popular consumer monitors with a commonly used research-grade accelerometer in a semistructured environment using a protocol that incorporates similar activities as those performed by adults on a daily basis.

## METHODS

Healthy adult men ( $n=15$ ) and women ( $n=15$ ) who were 18–80 years of age and able to participate in moderate-to-vigorous PA participated in this study. The study was approved by Ball State University's Institutional Review Board, and all subjects provided informed consent prior to participation. Age, height and weight of subjects were  $49.2\pm 19.2$  years,  $174.0\pm 8.9$  cm, and  $79.2\pm 15.5$  kg, respectively. Subjects were predominantly right-hand dominant (93.4%).

Four consumer monitors (FO, FZ, FF and JU), one research-grade accelerometer (AG) and a COSMED K4b2 (COSMED Srl, Rome, Italy) portable metabolic analyser were used in this study. All equipment was initialised with subject's sex, height, weight, and age, and synchronised to an external clock at the beginning of each visit.

The FO and FZ were mounted on a waist-worn elastic belt over the left hip, near the anterior axillary line, and were counterbalanced for anterior and posterior placement on the hip among subjects. The FF and JU were worn on the non-dominant wrist and counterbalanced for proximal and distal wrist placement among subjects. Two iPod Touch (Apple, Cupertino, California, USA) media players equipped with the Fitbit and Jawbone applications were synced to the FF and JU, and steps, EE, and active minutes were recorded. For the FO and FZ, steps and EE were recorded from the screen displays; active minutes were not assessed as these data are not available from the screen displays.

The AG (GT3X+), a commonly used accelerometer, was placed by research staff over participants' right hip on an elastic waistband at the anterior axillary line. AG data were recorded at a frequency of 60 Hz and analysed in 30 s epochs. All time with  $\geq 2691$  vector magnitude counts/min was used to estimate active minutes (measurement of active minutes (MVPA);  $\geq 3$  metabolic equivalents (METs)<sup>9</sup> The work-energy theorem and Freedson 2011 combination was used to calculate EE across the entire protocol from the AG. When selected, this combination automatically uses the work-energy theorem to calculate EE for sedentary and light PA, and the Freedson 2011 equation to calculate EE for MVPA.<sup>10</sup> All calculations were performed via ActiLife 6 software (ActiGraph).

The COSMED K4b2 was used to measure oxygen consumption ( $\text{VO}_2$ ) and carbon dioxide production during the study protocol. Breath-by-breath measurements were collected via a

breathing mask worn by participants and were used to determine  $\text{VO}_2$  in litres per minute (L/min), which was converted by a technician to EE by multiplying by 5 kcal per L of  $\text{O}_2$ .<sup>11</sup> All time with an EE  $\geq 3.0$  METS was summed for a measure of active minutes. The COSMED has been shown to provide accurate and reliable measures of  $\text{VO}_2$  over a wide range of activity intensities in comparison with metabolic carts and was used as the criterion measure for EE and active minutes in this study.<sup>12</sup>

Subjects participated in an 80 min, semistructured activity protocol, performing  $\geq 12$  activities from a list of 21 choices. Activities were grouped into the following categories: (1) sedentary activities (lying down, watching television, writing, reading, playing cards, and computer use), (2) household activities (standing, dusting, sweeping, vacuuming, folding laundry, making bed, picking up items from floor, and gardening) and (3) ambulatory and cycling activities (slow overground walk, brisk overground walk, treadmill walk, overground jog, treadmill jog, stair climbing, and stationary cycling). Subjects chose the pace, duration (2–15 min) and order of activities. At least four activities from each category (sedentary, household, and ambulatory) were performed, and subjects were instructed to spend  $\geq 40$  min in sedentary activities (to replicate adults' free-living sedentary behaviour patterns).<sup>12,13</sup> PA variables were recorded from the consumer monitors at the beginning and end of the protocol; therefore, activity-specific analyses could not be conducted for these data.

A trained research assistant counted and recorded steps during the entire protocol using a handheld tally counter; this served as the criterion measure of steps taken. A step was defined as lifting the entire foot and then placing it on the ground. During cycling, steps were counted for each pedal stroke, or two steps per revolution.

Repeated measures analysis of variance statistical tests were performed to assess differences from all four consumer monitors, AG, and criterion measures for steps, EE, and active minutes. This analysis was conducted for the total sample, as well as for a subgroup of the sample who did not perform cycling ( $n=9$ ). When the test statistic was significant, post hoc pairwise comparisons were performed using paired t-tests and a least significant difference correction. MAPE and per cent bias (%bias) were calculated to analyse predictive error of each PA monitor compared with the criterion measures; criterion measures included indirect calorimetry for EE and active minutes and counted steps by trained research technicians counting all steps taken during the protocol. Bland-Altman statistics were performed to determine the limits of agreement for each device compared with the criterion measure. Finally, a correlation analysis was performed among all devices, including the criteria.

## RESULTS

During the protocol, subjects averaged  $3231\pm 571$  steps, a total EE of  $294\pm 56$  kcal and  $28\pm 6$  active minutes. Compared with criterion measures, all monitors underestimated steps and EE (table 1;  $p<0.01$ ); additionally all monitors that assessed active minutes (FF, JU, and AG) underestimated this variable ( $p<0.05$ ). The FF had the lowest %bias (table 1) when counting steps and predicting EE, although this difference was only statistically significant when compared with the JU and AG ( $p<0.05$ ). The JU had the lowest %bias in predicting active minutes; however, this difference was only significantly different from the FF ( $p<0.05$ ). Similar PA underestimations were seen in the subgroup who did not cycle, although in some cases the differences were no longer statistically significant due to the small sample size (table 2).

**Table 1** Mean and %bias of PA variables measured by consumer monitors and research-grade accelerometer

PA monitor	Steps		Energy expenditure (kcal)		Active minutes	
	Mean±SD	%Bias	Mean±SD	%Bias	Mean±SD	%Bias
Total sample (n=30)						
Criterion	3231±571	n/a	294±56	n/a	28±6	n/a
FZ	2382±758*†	-26	217±57†	-26	n/a	n/a
FO	2388±760*†	-26	247±57*†	-16	n/a	n/a
FF	2486±748*†	-23	255±60*†	-13	11±8*†	-65
JU	2329±874†	-28	210±66*†	-29	19±7†	-35
AG	2183±821†	-32	217±69†	-26	17±8†	-40

\*Indicates significant difference from the AG.

†Indicates significant difference from criterion measure.

AG, ActiGraph GT3X+; FF, Fitbit Flex; FO, Fitbit One; FZ, Fitbit Zip; JU, Jawbone UP24; n/a, not applicable; PA, physical activity.

While all consumer monitors tested had higher accuracy than the AG for at least one PA variable, none of the consumer monitors had higher accuracy than the AG for all PA variables tested. The FO, FZ, and FF had significantly more steps compared with the AG (mean difference: 199–302 steps,  $p < 0.05$ ), whereas there was no significant difference in steps recorded by the JU compared with the AG (mean difference: 145 steps,  $p = 0.21$ ). The FZ produced similar estimates of EE compared with the AG; however, EE estimates were significantly higher for the FO and FF and lower for the JU (mean difference: -7 to 38 kcal). The FF recorded significantly fewer active minutes than the AG ( $p = 0.001$ ), while the JU similarly estimated active minutes compared with the AG.

Figure 1 reports MAPE for all PA monitors tested. The MAPE for the FF was significantly smaller than all other monitors ( $p < 0.05$ ), except the FZ when measuring steps and significantly smaller than the FO and JU when estimating EE ( $p < 0.05$ ). For active minutes, the JU and AG showed similar MAPE, both of which were significantly lower than the FF ( $p < 0.05$ ).

\* significantly different than actigraph accelerometer ( $p < 0.05$ )

The Bland-Altman plots for steps, EE, and active minutes are shown in figures 2A–E, 3A–E and 4A–C, respectively. There were wide limits of agreement for all devices and variables, indicating high individual predictive error. Additionally, the Bland-Altman plots show a trend of underestimation for all devices across all variables, with similar levels of error and variability for the consumer monitors as with the AG.

The correlation analysis for all devices is shown in table 3. For steps and EE, all monitors were significantly correlated with each other and with the criterion measures ( $p < 0.05$ ). The FF had poor, non-significant correlations with the criterion measure and the AG for active minutes. While not statistically evaluated,

the wrist-worn monitors (FF and JU) appeared more highly correlated with each other than the hip-worn monitors (FU, FZ, AG), and vice versa for the hip-worn monitors. The criterion measure correlations also appeared higher with the hip-worn monitors compared with the wrist-worn monitors.

## DISCUSSION

Our primary finding was that the consumer monitors underestimate all steps, EE, and active minutes to a similar degree as the AG, with no monitor consistently outperforming the others. Our findings extend previous research that assessed these PA variables individually, illustrating similar degree of error and good agreement in PA estimates produced by consumer monitors and the AG when measuring steps, EE, and active minutes.<sup>4–6 14</sup> A previous analysis by our research group indicated that, while sedentary and ambulatory (walking and jogging) activities could be measured accurately by consumer monitors, the consumer monitors tended to underestimate PA for activities in the household category and for cycling.<sup>15</sup>

All participants in the current study performed activities in the household category, but a subgroup analysis was performed for subjects who chose not to cycle in the protocol to better understand what activities contributed to underestimates of PA. Although the small sample ( $n = 9$ ) resulted in lack of statistical significance for some comparisons, there were still underestimations of PA variables by all monitors. Therefore, underestimations of overall PA found by all monitors in the current study were most likely driven by the periods of time during the protocol in which the subjects performed household activities, where shuffling of feet and slower walking speeds are common. For individuals with high sedentary time and/or ambulatory

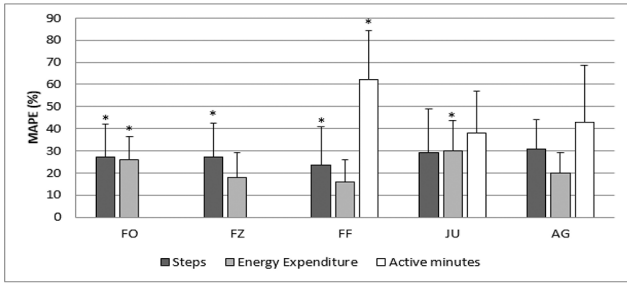
**Table 2** Mean and %bias of PA variables measured by consumer monitors and research-grade accelerometer for subjects who did not cycle

PA monitor	Steps		Energy expenditure (kcal)		Active minutes	
	Mean±SD	%Bias	Mean±SD	%Bias	Mean±SD	%Bias
Subgroup not cycling (n=9)						
Criterion	3399±585	n/a	304±47	n/a	29±8	n/a
FZ	2954±709*†	-13	282±47*	-12	n/a	n/a
FO	2935±717*†	-14	245±64†	-19	n/a	n/a
FF	2977±686†	-12	274±61*	-16	12±10*†	-62
JU	2838±897†	-17	241±70†	-25	23±7†	-26
AG	2822±685†	-17	226±37†	-26	25±6	-25

\*Indicates significant difference from the AG.

†Indicates significant difference from criterion measure.

AG, ActiGraph GT3X+; FF, Fitbit Flex; FO, Fitbit One; FZ, Fitbit Zip; JU, Jawbone UP24; n/a, not applicable; PA, physical activity.



**Figure 1** MAPE of physical activity monitors compared with criterion measures. FF, Fitbit Flex; FO, Fitbit One; FZ, Fitbit Zip; JU, Jawbone UP24; MAPE, mean absolute per cent error.

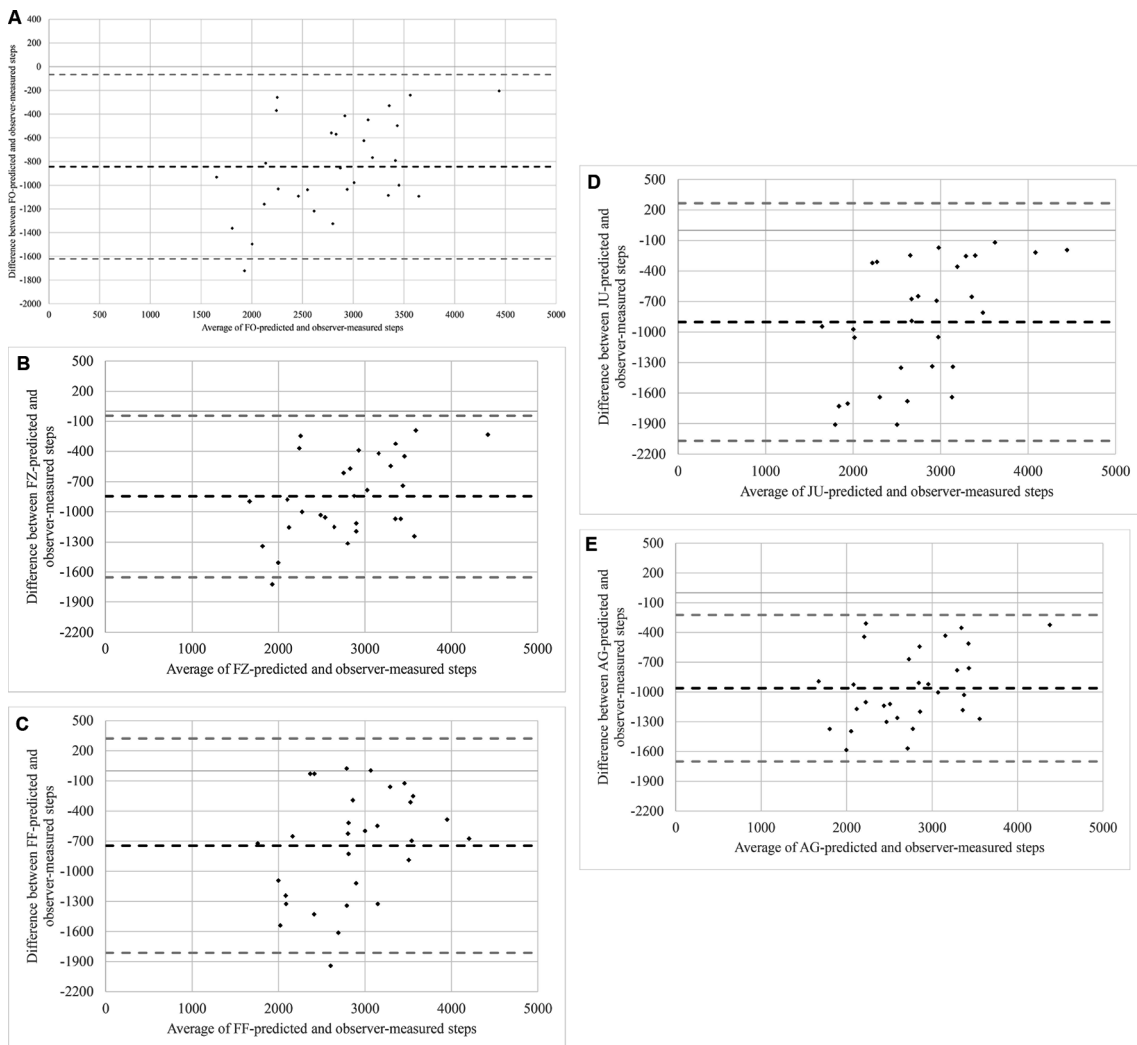
time throughout the day, these monitors will likely yield higher measurement accuracy than for individuals who spend more time in non-sedentary, non-ambulatory activities, such as cycling or household activities.

### The role of the research-grade accelerometer in our study

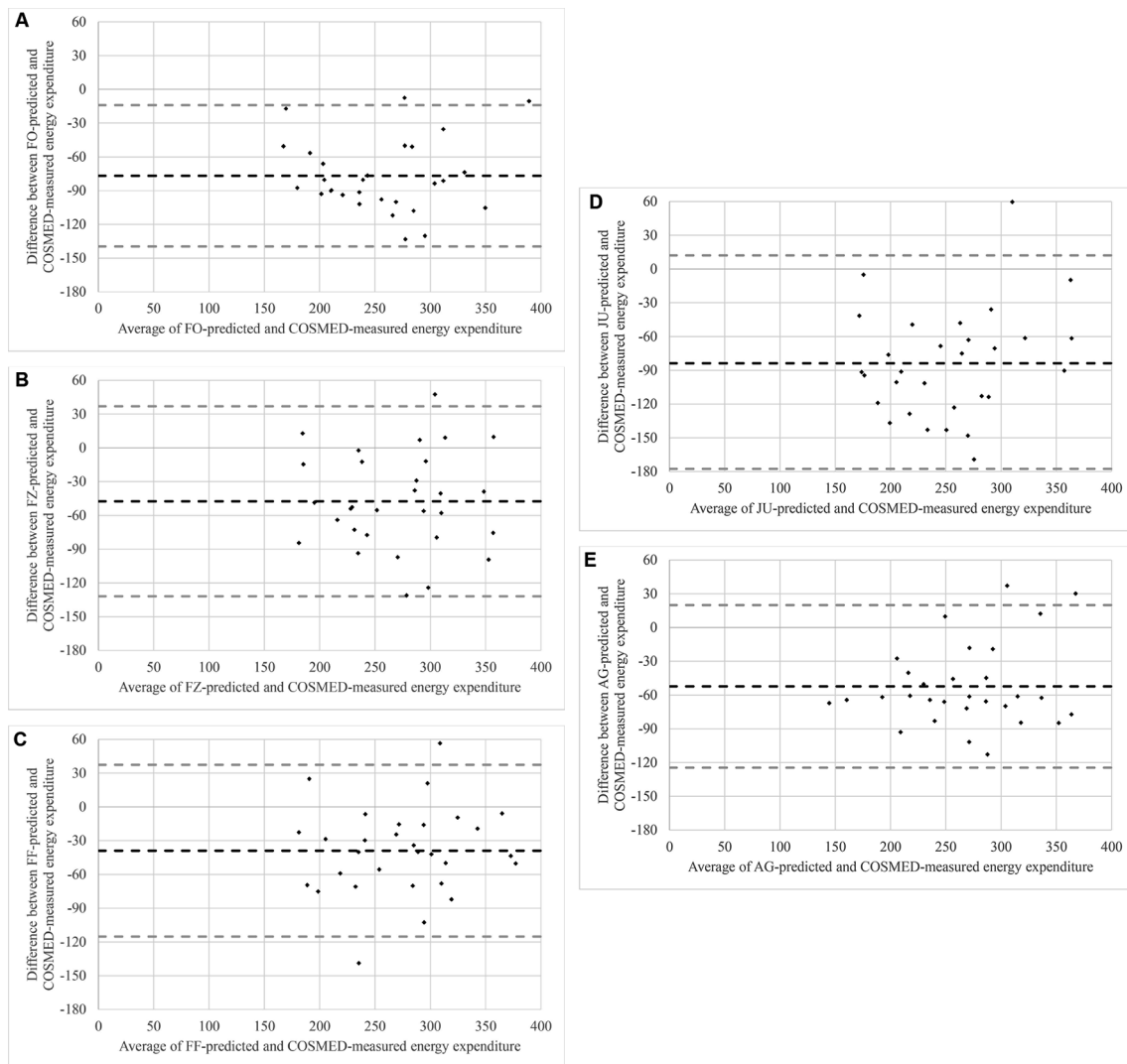
The inclusion of a research-grade accelerometer (AG) as a comparison in the current study provides a unique advantage of this study, as the AG is a popular research-grade accelerometer

that has been used extensively in various healthy and clinical populations, including the National Health and Nutrition Examination Survey,<sup>16</sup> to gain insight into their activity levels. Moreover, these devices have been used more commonly than any other accelerometer brand in intervention protocols to track adherence to prescribed PA.<sup>17</sup> Furthermore, the 2008 Physical Activity Guidelines for Americans recommends prescribing exercise using PA measures of intensity, active time and volume (as reported as MET-min/week), and these recommendations are based on measurements collected using accelerometry.<sup>1</sup> Thus, understanding how well consumer monitors measure PA compared with a commonly used research-grade accelerometer is important when comparing data collected from different consumer monitors.

Note that the AG, coupled with traditional (cut-point based) data analysis methods, demonstrated similar trends as the consumer monitors in underestimating PA variables in a simulated free-living environment, with the greatest error found in the MVPA. Fitbit's website states that active minutes must be accumulated in 10 min bouts to be recorded, similar to recommendations from PA guidelines; this feature of their proprietary software likely contributed to the underestimation of active minutes by all Fitbit monitors in this study.<sup>18</sup> Individuals considering using PA monitoring devices should be aware of this



**Figure 2** Bland-Altman plot for (A) Fitbit One (FO) steps, (B) Fitbit Zip (FZ) steps, (C) Fitbit Flex (FF) steps, (D) Jawbone UP24 (JU) steps and (E) ActiGraph GT3X (AG) compared with observer-measured steps.



**Figure 3** Bland-Altman plot for (A) Fitbit One (FO) energy expenditure, (B) Fitbit Zip (FZ) energy expenditure, (C) Fitbit Flex (FF) energy expenditure, (D) Jawbone UP24 (JU) energy expenditure and (E) ActiGraph GT3X+ (AG) energy expenditure compared with COSMED measured energy expenditure.

underestimation when tracking their progress and adherence to PA recommendations and/or exercise prescription. The use of PA logs to supplement objective data collection or the use of alternate exercise prescriptions based on daily step count should be considered to better capture PA, especially non-ambulatory types.

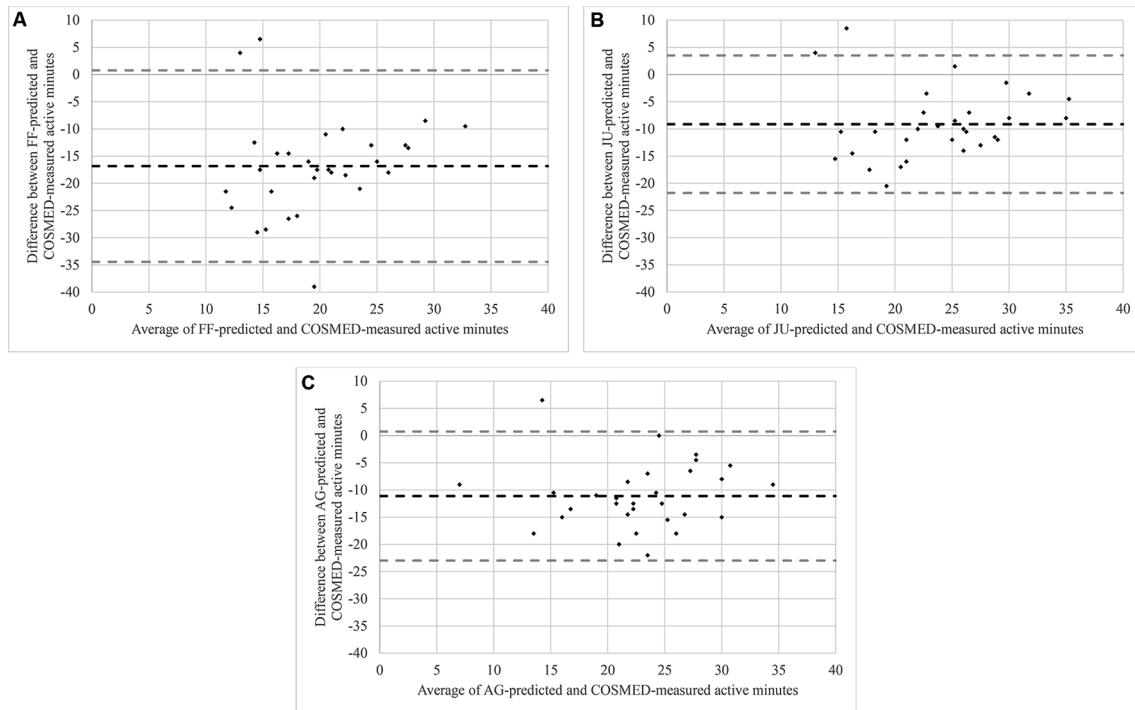
### Role of the monitor location (hip, wrist)

We were not surprised that the hip-worn monitors correlated more highly with each other than with the wrist-worn monitors (and vice versa) and also appeared to have higher correlations with the criterion measures. These findings are in agreement with past literature, which generally supports PA measurement accuracy with hip-worn compared with wrist-worn accelerometers.<sup>19–21</sup> However, wrist-worn devices are more popular considering that most consumer monitors are designed for wear on the wrist, and previous work also supports higher compliance with wrist-worn devices compared with hip-worn devices.<sup>22</sup> Therefore, choice of wrist-worn or hip-worn device will depend on the importance of compliance versus accuracy desired in a study.

### Research and clinical implications

Although the consumer monitors tested in this study had similar accuracy to the AG monitor and associated linear regression EE and activity intensity prediction, our findings should not be taken to mean that consumer monitors are on par with the highest accuracy monitoring devices available. For example, the activPAL, another popular research-grade accelerometer, has shown higher accuracy for measurement of sedentary behaviour and steps than the AG but poor for measurement of EE.<sup>23–25</sup> Additionally, as techniques for analysing research-grade accelerometer data improve (eg, through techniques such as pattern recognition), the AG research-grade accelerometer and other research-grade devices may become a more accurate method for assessing PA levels.

Our results indicate that a select sample of consumer monitors provided similar strengths and weaknesses and similar PA estimates to a single, popularly used research-grade device and associated prediction equation. Given the ease of use, relatively low cost, and comparable accuracy of consumer monitors to the AG in monitoring PA, consumer monitors may have utility in monitoring PA behaviours. Most consumer monitors also provide real-time feedback, which may influence behaviour



**Figure 4** Bland-Altman plot for (A) Fitbit Flex (FF) predicted active minutes, (B) Jawbone UP24 (JU) predicted active minutes and (C) ActiGraph GT3X+ (AG) predicted active minutes compared with COSMED measured active minutes.

via the Hawthorne effect or a better awareness of PA patterns. Therefore, consumer monitors may be less appropriate for use in surveillance studies or to assess the effectiveness of interventions because of their potential to influence PA behaviour more than the AG or other research-grade accelerometers that do not give immediate feedback. On the contrary, their real-time feedback may serve as a valuable intervention strategy/motivational tool

that could help promote adoption/maintenance of healthy PA habits in the general population; if used for such purposes, the accuracy of the monitors would be of less importance than the fact that they may encourage people to be more active. However, individuals should be cautious when comparing PA values from different consumer monitors due to discrepancies seen in this and previous studies.<sup>4-6</sup> Additionally, individuals should consider these monitors' potential underestimation of PA variables when provided feedback from the monitors and consider if other types of monitors (such as the activPAL) may give more accurate data to meet their specific needs. Thus, the decision to use consumer monitors or research-grade monitors will depend on the goals of the assessment, the options available to the researcher or individual and the relative importance of usability versus accuracy of measurement.<sup>2</sup>

### Limitations

This study comprised a range of activities performed during a semistructured protocol, but had a small sample size. Given the slower movement speeds with older, diseased or disabled adults and poorer accuracy of PA monitors for measuring slower movement speeds,<sup>26</sup> measurement error may be greater in these populations than indicated by our results. Additionally, although similar activities were performed in the protocol as those commonly performed on a daily basis, the study did not use a true free-living setting. Measurement error may differ in a true free-living environment than indicated by the current protocol, especially if the proportion of the day spent in different types of activities differs from that performed in the current study, where participants spent roughly 50-60% of the time in sedentary activities<sup>27</sup>. This study only assessed the total measurement values for steps, EE, and active minutes over the observed period for all monitors assessed, and therefore time-matched analyses for each activity performed are not provided in current results. However, this analysis can be found in a similar study

**Table 3** Correlations among monitors and criterion measures for steps, energy expenditure and active minutes

Steps	Criterion	FZ	FO	FF	JU	AG
Criterion		0.86	0.85	0.69	0.74	0.85
FZ			1.00	0.78	0.85	0.97
FO				0.77	0.83	0.97
FF					0.92	0.79
JU						0.83
AG						
Energy expenditure	Criterion	FZ	FO	FF	JU	AG
Criterion		0.84	0.71	0.78	0.69	0.81
FZ			0.82	0.83	0.88	0.78
FO				0.82	0.89	0.84
FF					0.90	0.80
JU						0.76
AG						
Active minutes	Criterion	FF	JU	AG		
Criterion		0.18*	0.54	0.57		
FF			0.46	0.10*		
JU				0.61		
AG						

\*Indicates that the two measures are **not** significantly correlated ( $p>0.05$ ).

AG, ActiGraph GT3X+; FF, Fitbit Flex; FO, Fitbit One; FZ, Fitbit Zip; JU, Jawbone UP24.

## What are the findings?

- ▶ All PA monitors tested (consumer-based and research-grade) tended to underestimate PA measurements compared with criterion methods.
- ▶ The hip-worn monitors correlated more highly with each other than with the wrist-worn monitors (and vice versa) and also appeared to have higher correlations with the criterion measures.

from our laboratory.<sup>15</sup> The current study defined a step as lifting the entire foot and then placing it on the ground, which may not be the same method as the PA monitors use to recognise a step. This may have been a cause of the underestimation by monitors in measuring steps specifically. Finally, although a subgroup analysis was performed for individuals who did not cycle in the protocol, the sample size was only nine participants. Further validation and comparison of these PA monitors is warranted to determine which monitor has the best capability of capturing non-ambulatory measurement, household chores, cycling-type exercise (ie, stationary bicycle, recumbent trainer) and sport activities.

## Summary and conclusion

In conclusion, consumer monitors and the AG research-grade accelerometer underestimated PA variables and EE assessed in comparison with criterion measures to a similar degree. Due to the underestimates of PA and discrepancies between consumer monitors, researchers and consumers should be cautious when comparing PA values that were obtained using different monitors.

## How might it impact on clinical practice in the future?

- ▶ Due to the differences found in the measures of PA between the consumer-based PA monitors and the ActiGraph during a simulated free-living protocol, one should be cautious when comparing PA values obtained using these monitors in free-living settings.
- ▶ Given the ease of use, relatively low cost, immediate feedback capabilities and comparable accuracy of consumer monitors with the AG in monitoring PA, consumer monitors may have utility in monitoring PA behaviours and promoting adoption/maintenance of healthy PA habits in the general population.
- ▶ This study shows hip-worn devices to have higher accuracy compared with wrist-worn devices; however, past literature has shown higher compliance using wrist-worn devices. Therefore, the choice of wrist-worn or hip-worn devices will depend on the importance of compliance versus accuracy.

**Funding** This study received funding support, in part, from the Ball State University ASPIRegrant.

**Competing interests** None declared.

**Patient consent** Obtained.

**Ethics approval** Ball State University Institutional Review Board.

**Provenance and peer review** Not commissioned; externally peer reviewed.

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

## REFERENCES

- 1 Physical Activity Guidelines Advisory Committee. *Physical activity guidelines Advisory Committee Report, 2008*. Washington, DC: U.S. Department of Health and Human Services, 2008.
- 2 Strath SJ, Kaminsky LA, Ainsworth BE, *et al*. Guide to the assessment of physical activity: clinical and research applications: a scientific statement from the American Heart Association. *Circulation* 2013;128:2259–79.
- 3 Dolan L. Fitbit, Jawbone, Nike had 97 percent of fitness tracker retail sales in 2013. <http://mobihealthnews.com/28825/fitbit-jawbone-nike-had-97-percent-of-fitness-tracker-retail-sales-in-2013/> (January 15, 2014).
- 4 Lee JM, Kim Y, Welk GJ. Validity of consumer-based physical activity monitors. *Med Sci Sports Exerc* 2014;46:1840–8.
- 5 Bai Y, Welk GJ, Nam YH, *et al*. Comparison of consumer and research monitors under semistructured settings. *Med Sci Sports Exerc* 2016;48:151–8.
- 6 Murakami H, Kawakami R, Nakae S, *et al*. Accuracy of wearable devices for estimating total energy expenditure: comparison with metabolic chamber and doubly labeled water method. *JAMA Intern Med* 2016;176:702–3.
- 7 Ferguson T, Rowlands AV, Olds T, *et al*. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. *Int J Behav Nutr Phys Act* 2015;12:42.
- 8 Tully MA, McBride C, Heron L, *et al*. The validation of Fitbit Zip™ physical activity monitor as a measure of free-living physical activity. *BMC Res Notes* 2014;7:952.
- 9 Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. *Med Sci Sports Exerc* 1998;30:777–81.
- 10 Sasaki JE, John D, Freedson PS. Validation and comparison of ActiGraph activity monitors. *J Sci Med Sport* 2011;14:411–6.
- 11 Pescatello LS, Arena R, Riebe D, *et al*; *ACSM's guidelines for exercise testing and prescription*. 9th edn. Philadelphia, PA: Lippincott Williams & Wilkins, 2014:88–93.
- 12 McLaughlin JE, King GA, Howley ET, *et al*. Validation of the COSMED K4 b2 portable metabolic system. *Int J Sports Med* 2001;22:280–4.
- 13 Matthews CE, Chen KY, Freedson PS, *et al*. Amount of time spent in sedentary behaviors in the United States, 2003–2004. *Am J Epidemiol* 2008;167:875–81.
- 14 Case MA, Burwick HA, Volpp KG, *et al*. Accuracy of smartphone applications and wearable devices for tracking physical activity data. *JAMA* 2015;313:625–6.
- 15 Nelson MB, Kaminsky LA, Dickin DC, *et al*. Validity of consumer-based physical activity monitors for specific activity types. *Med Sci Sports Exerc* 2016;48:1619–28.
- 16 Loprinzi PD, Lee H, Cardinal BJ, *et al*. The relationship of actigraph accelerometer cut-points for estimating physical activity with selected health outcomes: results from NHANES 2003–06. *Res Q Exerc Sport* 2012;83:422–30.
- 17 Montoye AH, Moore RW, Bowles HR, *et al*. Reporting accelerometer methods in physical activity intervention studies: a systematic review and recommendations for authors. *Br J Sports Med* 2016.
- 18 Fitbit Help: Fitbit. [https://help.fitbit.com/articles/en\\_US/Help\\_article/1379](https://help.fitbit.com/articles/en_US/Help_article/1379) (May, 2016).
- 19 Ellis K, Kerr J, Godbole S, *et al*. Hip and wrist accelerometer algorithms for free-living behavior classification. *Med Sci Sports Exerc* 2016;48:933–40.
- 20 Ellis K, Kerr J, Godbole S, *et al*. A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers. *Physiol Meas* 2014;35:2191–203.
- 21 Swartz AM, Strath SJ, Bassett DR, *et al*. Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Med Sci Sports Exerc* 2000;32(9 Suppl):S450–S456.
- 22 Troiano RP, McClain JJ, Brychta RJ, *et al*. Evolution of accelerometer methods for physical activity research. *Br J Sports Med* 2014;48:1019–23.
- 23 Kozey-Keadle S, Libertine A, Lyden K, *et al*. Validation of wearable monitors for assessing sedentary behavior. *Med Sci Sports Exerc* 2011;43:1561–7.
- 24 Powell C, Carson BP, Dowd KP, *et al*. Simultaneous validation of five activity monitors for use in adult populations. *Scand J Med Sci Sports* 2016.
- 25 Lyden K, Kozey Keadle SL, Staudenmayer JW, *et al*. Validity of two wearable monitors to estimate breaks from sedentary time. *Med Sci Sports Exerc* 2012;44:2243–52.
- 26 Phillips LJ, Petroski GF, Markis NE. A comparison of accelerometer accuracy in older adults. *Res Gerontol Nurs* 2015;8:213–9.
- 27 Donaldson SC, Montoye AH, Tuttle MS, *et al*. Variability of objectively measured sedentary behavior. *Med Sci Sports Exerc* 2016;48:755–61.