Validation of Accelerometer-Based Estimations of Energy Expenditure During High-Intensity Interval Training

Nicholas Remillard, Marisa Mulvey, Gregory Petrucci Jr, &John R Sirard**

University of Massachusetts Amherst, 30 Eastman Lane, Amherst MA 01003 and 240 Thatcher Road, Amherst, MA 01003

https://doi.org/10.33697/ajur.2022.057

Students: nremillard@umass.edu, mmulvey@ufl.edu, gpetrucci@umass.edu*
Mentor: jsirard@kin.umass.edu

ABSTRACT
Accelerometers are used to assess free-living physical activity (PA) and energy expenditure (EE). Energy expenditure estimation algorithms have been calibrated using steady-state exercise. However, most free-living PA is not steady-state. **Objective:** The purpose of this study was to discern the differences between criterion-measured and accelerometer-estimated EE (kCals) during a non-steady-state High-Intensity Interval Training (HIIT) session. **Methods:** Recreationally active adults (N=29, 18-30 years) completed one of two HIIT protocols. Each participant wore ActiGraph GT3X+ accelerometers on the right hip and non-dominant wrist while EE was measured using portable indirect calorimetry. Data analysis was conducted using custom R scripts and bias [95% CIs] to determine significant differences between indirect calorimetry and EE estimates using previously developed algorithms. **Results:** All accelerometer algorithms underestimated EE during recovery intervals (range; -4.31 to -6.55 kCals) and overestimated EE during work intervals (0.57 to 5.70 kcals). Over the whole HIIT session, only the Hildebrand wrist method was not significantly different from the criterion measured EE. **Conclusion:** Current ActiGraph EE estimations based on steady-state activities underestimate EE during recovery periods of treadmill HIIT sessions. Future studies should investigate accelerometer signals immediately after high-intensity bouts to more accurately predict EE of the subsequent recovery period.

KEYWORDS
ActiGraph; Accelerometer; HIIT; Indirect calorimetry; EPOC; Energy expenditure; Non-steady state; Calories

INTRODUCTION

Accelerometers have been used for over two decades to estimate energy expenditure (EE) and, from these estimates, categorize physical activity (PA) intensity levels, as defined by the US PA Guidelines.[1,2] Accelerometers have become a standard epidemiological tool to objectively measure physical activity and relate minutes spent in sedentary time or moderate-to-vigorous physical activity with other health behaviors and outcomes.[3–5] However, the inability of existing algorithms to estimate EE during non-steady-state activities is a known limitation.

By calibrating accelerometer EE estimation algorithms only on steady-state aerobic exercise results in a relatively narrow representation of how people exercise and perform most physical activities. Currently, no accelerometer-estimated EE algorithm considers the lag in aerobic metabolism at the beginning of a bout of exercise (oxygen deficit) or the elevated metabolism that occurs after an exercise bout (excess post-exercise oxygen consumption; EPOC). From their inception, laboratory-based accelerometer calibration studies have consistently removed the first and last minute of an exercise bout in order to only model EE from the steady-state part of the activity. Therefore, accelerometers consistently underestimate overall EE outside of structured steady-state exercise bouts, supported by one study that identified elevations in EE for 14 hours following a vigorous bout of exercise, accounting for an additional 190+71.4 kCals occurring after the exercise bout.[6-9] This accelerometer algorithm limitation is especially concerning since interval training (repeated short-duration high-intensity intervals separated by low-intensity intervals) and other forms of non-steady-state exercise have gained popularity among healthy and clinical populations.[10]

High-intensity interval training (HIIT) is broadly defined as brief (6s-4min), high intensity (85-250% VO_2 max or 90-100% HR_{max}) exercise bouts separated by similar or longer (10s-5min) recovery bouts of lower intensity (ranging from passive recovery to active recovery between 20-40% VO_2 max).[10,11] A variety of HIIT protocols have become widely accepted as an alternative form of exercise to continuous steady-state aerobic exercise, providing similar benefits, including reduced cardiometabolic disease risk factors.[10,12–16] For this initial research study, a structured HIIT session fitting the above definition of a HIIT session and including an exercise modality consistent with previous accelerometer calibration studies was used as the experimental paradigm to assess the error of accelerometer-based EE estimates during intermittent activity. Focusing this research on time periods of expected error, like EPOC, will lead to evidence that may allow for improved estimation of EE using wearable devices. Therefore, the purpose of this proof-of-concept study was to describe the discrepancies in EE estimated from hip- and wrist-worn accelerometers during two different HIIT protocols compared with indirect calorimetry (criterion measure of EE). We hypothesized that 1) during the work intervals, hip accelerometer estimates of EE would be similar to indirect calorimetry while wrist estimates of EE would overestimate EE, 2) during recovery intervals, all accelerometer estimates would underestimate EE, resulting in 3) underestimations of total EE for all accelerometer methods over the whole HIIT session.

METHODS AND PROCEDURES

Research Design

All participants completed a preliminary and a HIIT session at least 24 hours apart and within two weeks of one another. The preliminary session was used to identify each participant's treadmill speed corresponding to 95% of their age-predicted maximal heart rate, which was used during the HIIT session for the work intervals. Two different HIIT protocols consistent with previous HIIT definitions, representing different work-to-recovery ratios (1:2 and 1:1), were conducted in independent samples.[10,11] The participants performing the first protocol completed the study prior to the larger sample of participants performing the second protocol. Protocol 1 was originally designed as a pilot study and is included in this manuscript with protocol 2 to emphasize the consistency in results. Protocol 2 improved upon protocol 1 by including an even gender distribution in a larger sample, and the different work-to-recovery ratios were included to observe potential differences in results among two common HIIT paradigms.

Participants

All participants were recruited on the University of Massachusetts Amherst campus via posted flyers, email outreach, and word of mouth. Inclusion criteria for both protocols (Protocol 1; N=9, 8 M, Protocol 2; N=20, 10 M) were 18-30 years, with a normal Body Mass Index (BMI; 18.0-24.9 $kg \cdot m^{-2}$) and physically active as defined by the Godin-Shephard Leisure-time Exercise Questionnaire, scoring 24 points or above in moderate or strenuous exercise and reported performing strenuous exercise at least once per week.[17] No participant possessed previous health conditions that would preclude their ability to participate in vigorous activity (assessed using the Physical Activity Readiness-Questionnaire [PAR-Q]).[18] In addition to the PAR-Q, participants also completed an informed consent form before the preliminary session. All study procedures were approved by the University's Institutional Review Board.

Indirect Calorimetry

The Oxycon Mobile (CareFusion Solutions, LLC, San Diego, CA) indirect calorimetry system provided the criterion measure of EE. The Oxycon Mobile is a portable gas-exchange analyzer, worn like a backpack, which includes a face mask covering the nose and mouth, sampling tubes attached to a gas analyzer, and a telemetry unit. Data from the telemetry unit was sent to a PC and processed in JLAB (CareFusion, Germany 234 Gmbh) software. The Oxycon gas analyzer was calibrated within a 3% difference with a known gas mixture of 15.98% O_2 and 3.90% CO_2, and volume calibration was completed through JLAB. Breath-by-breath O_2 and CO_2 data were collected and averaged every 5 seconds in units of mL/min.

Heart Rate

Heart rate (HR) was measured during the preliminary session using a Polar monitor (Polar Electro Oy, Kempele, Finland) strapped to the participant's chest. The research staff recorded HR values per the study protocol outlined below.

Accelerometers

Two ActiGraph GT3X+ accelerometers (ActiGraph Corp, Pensacola, FL), one on the right hip and the other on the non-dominant wrist, were used in both protocols. The ActiGraph was chosen because it is the most common accelerometer in United States' epidemiological studies measuring physical activity.[5] These locations were chosen because these are the most common wear locations for wearable accelerometer devices, and the accelerometer algorithms chosen for this study were calibrated using the right hip and non-dominant wrist locations. The accelerometers measured accelerations along three axes at a sampling rate of 80 Hz, and data were then collapsed to 1-second epochs. ActiLife 6 software was used to initialize and download devices.

Preliminary session

The goal of each preliminary session was to determine the treadmill velocity at which 95% of the maximal heart rate (HR_{max}) would be achieved, representing a vigorous intensity that fits the work interval criteria of HIIT exercise.[10] The preliminary session began with a 5-minute warm-up walking period at 4.8 km/hr. Then, participants were instructed to run at 8.9, 9.7, and 10.5 km/hr for two minutes each. Three HR data points were collected at each of the three velocities (at 40 seconds, 80 seconds, and 120 seconds), and the average of the three values was used for further calculations. After the final running bout, participants were instructed to engage in a cool-down period for 5-minutes at 4.8 km/hr. A linear fit of HR over treadmill speed was used to estimate the speed at which the participant would attain 95% of their age-predicted HR_{max} (HR_{max}=220-age).[19,20]

HIIT session
Participants were required to fast for 4 hours prior to the HIIT session to ensure that the EE predictions were attributed to physical activity energy expenditure (PAEE) rather than diet-induced thermogenesis. The HIIT session began with a 5-minute warm-up at 4.8 km/hr, followed by the HIIT intervals when participants alternated between five work intervals and five recovery intervals. Participants ran at the calculated velocity from the preliminary session during work intervals and walked at 4.8 km/hr during the recovery intervals. Protocol 1 (N=9) consisted of 45-second work intervals and 90-second recovery intervals (1:2 work-to-recovery ratio), totaling 11 minutes and 15 seconds of interval training. Protocol 2 (N=20) consisted of 1-minute work intervals and 1-minute recovery intervals (1:1 work-to-recovery ratio), totaling 10 minutes of interval training. All interval sessions were followed by a 5-minute walking cool-down at 4.8 km/hr (see Figure 1).

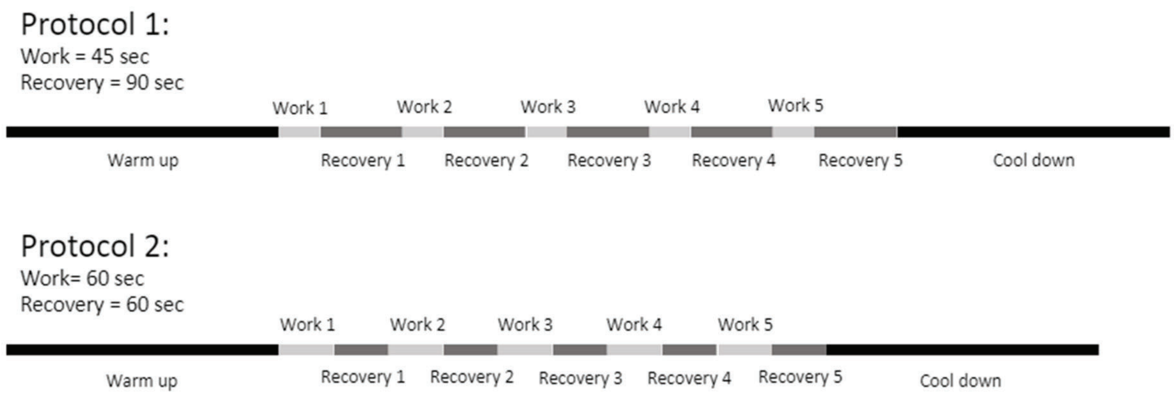

Figure 1. Sequence of events for Protocol 1 (1:2) and Protocol 2 (1:1)

Data Processing
The Oxycon Mobile collected breath-by-breath data during the entire exercise session, including warm-up and cool-down, and was reported in 5-second epochs, which represented the rate of oxygen consumption, in mL/min, for each 5-second epoch. Only data between the start of the first work interval to the end of the last rest interval was extracted for analysis. The average of twelve data points (epochs) provided the average oxygen consumption per minute. Minute-by-minute oxygen consumption data was converted to EE, in kCals/minute, using *the Weir Equation:* [21]

$$\text{EE (kCals/minute)} = [3.94*(\text{VO}_2 \text{ in L/min})]+[1.11*(\text{VCO}_2 \text{ in L/min})]$$

Four common accelerometer-based EE prediction algorithms were compared with our criterion measure, each using different accelerometer processing or location. To match the 5-second metabolic data, ActiGraph data (counts) were summed for each 5-second epoch, synchronous with the Oxycon Mobile data. The ActiGraph counts per 5-second epoch were then summed to create counts per minute data by summing twelve consecutive 5-second epochs. Vertical count data from the hip-worn ActiGraph were converted to relative units of energy expenditure using the *Freedson Equation*, where BW (body weight) is measured in kilograms (kg):[1]

$$\text{EE (kCals/minute)} = [(0.00094)(\text{counts/min})]+[(0.1346)(\text{BW})]-(7.37418)$$

The ActiGraph hip VM (vector magnitude; incorporating all three axes) count data were converted with the *Sasaki Equation*, producing EE predictions expressed in METs per minute:[22]

$$\text{EE (METs/minute)} = (0.000863*\text{VM})+0.668876$$

METs per minute were then converted into kCals per minute using the following equation, where BW is measured in kg (https://sites.google.com/site/compendiumofphysicalactivities/help/unit-conversions):

$$\text{EE (kCals/minute)} = (\text{METs}*3.5*\text{BW})/200$$

In addition to equations using count data, raw acceleration data from the hip and wrist were translated to Euclidian Norm Minus One (ENMO) units (R-code used to calculate ENMO can be found in Hildebrand et al., 2014) and then to VO_2 using the Hildebrand hip and wrist equations.[23] Both equations produce VO_2 (mLO_2/kg/min) using the ENMO data (m*g* = ENMO*1000):

$$\text{Hip: } VO_2\ (mLO_2/kg/min) = 0.0554*(mg)+6.67$$
$$\text{Wrist: } VO_2\ (mLO_2/kg/min) = 0.0320*(mg)+7.28$$

Resulting VO_2 from the Hildebrand equations was converted to kCals per minute using the Weir equation, including only VO_2:

$$\text{kCals/minute} = 5*(VO_2 \text{ in L } O_2/\text{min})*BW$$

Data Analysis

Software-specific output files from the Oxycon Mobile and ActiLife were converted to CSV files, and custom R scripts were used for all data processing and analyses. No data from the warm-up or cool-down periods are included in the analyses. Analyses were performed separately for each protocol since they were not designed to be iso-caloric. ActiGraph-estimated EE has been plotted along with the measured EE calculated from the oxygen consumption data to visualize EE patterns during the HIIT sessions. Descriptive analyses indicated that the data approximated a normal distribution for all statistical tests.

To assess the criterion validity of accelerometer-based EE estimates during the HIIT protocols, the difference (bias) between ActiGraph-estimated EE and indirect calorimetry was calculated. Separate bias estimates were calculated for the differences in EE across the whole HIIT session, just the work intervals and just the recovery intervals. Mean bias and 95% bias intervals that did not span zero defined significant differences between ActiGraph-estimated and indirect calorimetry measured EE. Bias values above zero indicate an ActiGraph overestimation, and bias values below zero indicate an Actigraph underestimation.

RESULTS

Based on the participant inclusion criteria, the samples for both protocols were similar in mean age and range (18 to 30 years for both samples) with normal BMI values and considered regularly active. Protocol 1 consisted of mostly males, while Protocol 2 had an even split between male and female participants (Table 1). All analyses in this paper were conducted, including both male and female participants together. There were observed absolute differences between the sexes, but the sex-specific bias estimates were similar to the full sample, so sex-separate analyses are not included here.

	Weight (kg)	**Height (m)**	**BMI (kg/m²)**	**Age (yrs)**	**Godin-Shepard Score (units)**	**Treadmill Work Speed (kmh)**
Protocol 1 (N=9, 8 male)	75.5 $\pm$ 10.9	1.7 $\pm$ 0.1	24.7 $\pm$ 3.7	20.4 $\pm$ 1.7	71.7 $\pm$ 16.6	12.8 $\pm$ 1.9
Protocol 2 (N=20, 10 male)	70.0 $\pm$ 7.7	1.7 $\pm$ 0.1	24.4 $\pm$ 2.9	20.7 $\pm$ 1.9	65.5 $\pm$ 14.3	12.1 $\pm$ 2.0

Table 1. Participant Characteristics (Mean $\pm$ SD).

Figure 2 provides a visualization of the ActiGraph and oxygen consumption data during the full duration of the HIIT sessions for Protocol 1 (panels A and C) and Protocol 2 (panels B and D). The accelerometer data clearly demonstrates an instantaneous response to movement changes, while the metabolic data show a lagged increase in O_2 consumption at the beginning of each work interval and a delay in the decline in O_2 consumption during the recovery intervals (demonstrating the effects of EPOC).

Figure 3 provides mean kCals, by method, for each work and rest interval for both protocols. Error bars are not presented to ease the readability of the figure. Supplemental Table 1 (Table S1) contains means and standard deviations for each work and recovery bout for both protocols. Consistent with the unprocessed data in Figure 1, EE measured with indirect calorimetry is greater during the recovery intervals due to EPOC compared with EE values during the work intervals. The difference in work-to-recovery ratios between protocols (protocol 1, 1:2; protocol 2, 1:1) explains the significantly higher EE values during Protocol 1 recovery intervals as compared to Protocol 2 recovery intervals ($p < 0.001$; two-sample heteroscedastic T-test to compare criterion measured EE between protocols).

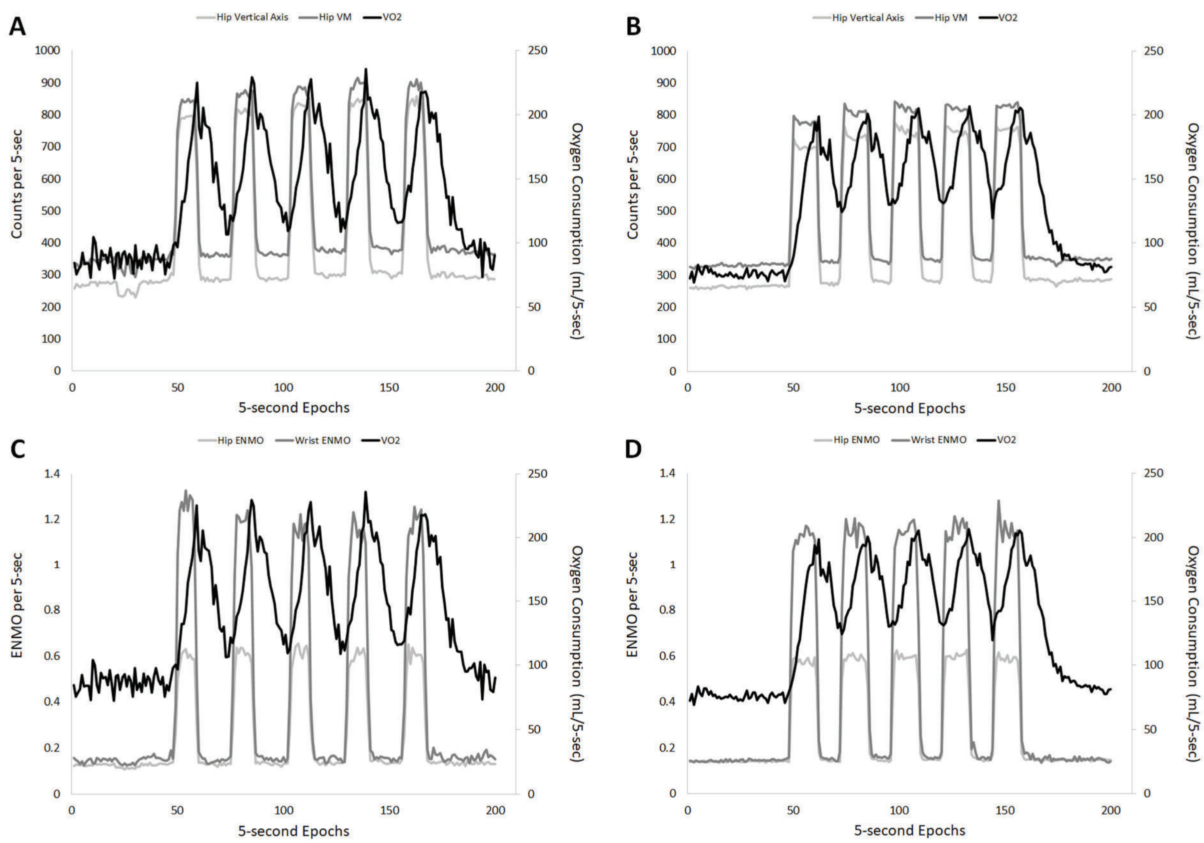

Figure 2. Mean accelerometer counts/5-sec interval for the hip vertical axis and hip VM (primary axis), as well as the rate of oxygen consumption (secondary axis; mL O_2/5-sec), are shown in Panel A (protocol 1) and Panel B (protocol 2). Hip and wrist ENMO data per 5-sec interval (primary axis), as well as the rate of oxygen consumption (secondary axis; mL O_2/5-sec), are shown in Panel C (protocol 1) and Panel D (protocol 2).

Total EE (kCals; mean (SD)) measured by the criterion measure for the entire HIIT session (calculated from the beginning of Work 1 through the end of Recovery 5) for Protocol 1 and Protocol 2 were 107.8 (19.5) kCals and 98.9 (24.7) kCals, respectively. Bias and 95% confidence interval estimates for each method, for the whole HIIT session (total), and separated into work and recovery intervals are presented in Table 2. Graphical representations of bias for each method can be found in Supplemental Figures S1-S4.

The Freedson algorithm significantly overestimated work EE for Protocol 1 (bias in kCals, 1.69), but not Protocol 2 (0.57); underestimated recovery EE for both protocols (-4.66 and –4.41); and therefore, underestimated total EE for both protocols (-14.82 and –19.25). For both protocols, the Sasaki and Hildebrand hip algorithms both significantly overestimated work EE (range, 1.30 to 4.00) and underestimated recovery EE (range, -4.31 to –5.63). The Sasaki algorithm underestimated total EE for both protocols (-11.77 and –15.08), while the underestimation of total EE using the Hildebrand hip algorithm was significant for protocol 1 (-8.15) but not for protocol 2 (-1.83). The Hildebrand wrist algorithm produced the greatest work EE overestimations (5.00 and 5.70) and the greatest recovery EE underestimations (-6.55 and -5.14). The Hildebrand wrist under- and overestimations canceled each other out such that there was no significant difference in total EE for either protocol (-7.76 and 2.79) compared with indirect calorimetry.

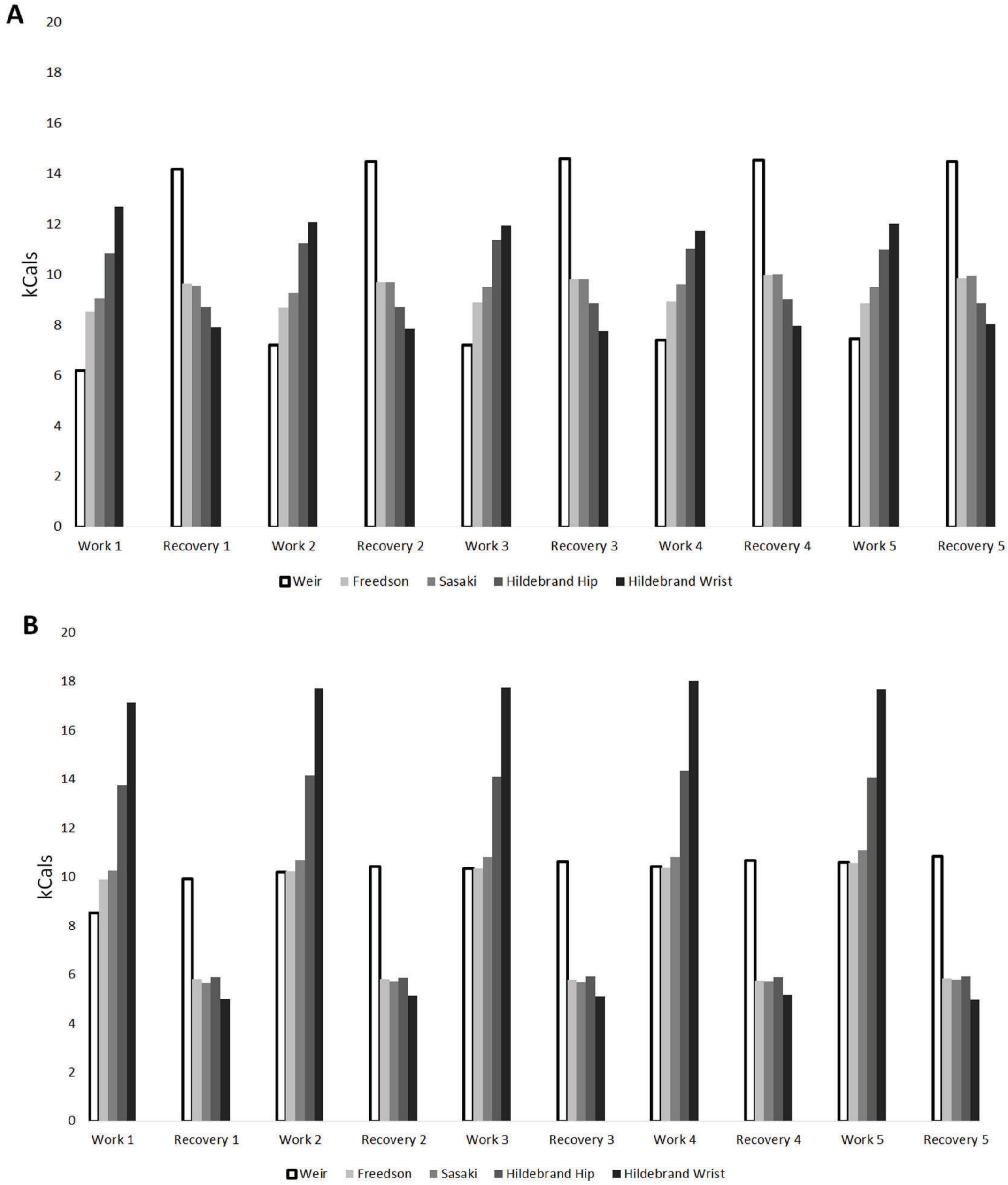

Figure 3. Energy expenditure (EE) in kCals/interval calculated from oxygen consumption data (criterion using the Weir equation) and the accelerometer prediction equations, Freedson Hip, Sasaki Hip, Hildebrand Hip, and Hildebrand Wrist for protocol 1 (panel A) and protocol 2 (panel B).

DISCUSSION

The purpose of this study was to characterize the discrepancies between accelerometer-estimated EE and indirect calorimetry during two HIIT protocols. Overall, most ActiGraph EE equations overestimated EE during work intervals and underestimated it during recovery intervals. Considering the total EE over the whole protocols, the Freedson equation underestimated EE for both protocols, the Hildebrand wrist equation was not biased for either protocol, and the Sasaki and Hildebrand hip equations underestimated for one protocol and were not biased for the other. These results have implications for the research community and for individuals wearing commercial-grade accelerometer devices.

To our knowledge, no other study has compared indirect calorimetry with accelerometer-estimated EE during the separate work and recovery intervals of intermittent exercise. However, the current findings are supported by previous free-living accelerometer validation studies that have tested existing algorithms to estimate EE using indirect calorimetry.[24–26] These previous studies have found that, for lifestyle activities like racquetball, basketball, and tennis, which involve intermittent movement, the Freedson prediction equation underestimated kCals as compared with indirect calorimetry. Underestimation during these studies ranged from –1.0 kCals/min to –5.5 kCals/min. In contrast, the validation studies considering steady-state ambulatory activities found less bias (between -0.8 and 1.8 kCals/min) using the Freedson equation, suggesting that existing energy expenditure estimation equations may be fairly accurate during steady-state treadmill activities.[24,26] Despite differences in protocols across studies, the underestimation of EE found during these intermittent leisure-time sports activities and ADLs support our findings from the current HIIT protocols and further indicates the need to develop more specific EE algorithms targeting intermittent activities. Because these previous studies only assessed the simulated free-living activity sessions as a whole, it remains unclear what during the session was driving the underestimation of EE during these studies.[24–26] The current study is novel in using the HIIT paradigm as a substitute for intermittent lifestyle activities and breaking up the activity period into work and recovery intervals to determine if undetected elevated EE during recovery intervals is responsible for accelerometer EE biases.

Kcals/min	Method	Protocol 1 Bias [95% C.I.]	Protocol 2 Bias [95% C.I.]
Total	Freedson	-1.32 [-2.41, -0.23] *	-1.92 [-2.83, -1.02] *
	Sasaki	-1.05 [-2.36, 0.26]	-1.51 [-2.44, -0.58] *
	Hildebrand Hip	-0.72 [-1.42, -0.02] *	-0.18 [-0.88, 0.51]
	Hildebrand Wrist	-0.69 [-1.64, 0.26]	0.28 [-0.64, 1.20]
Work	Freedson	2.26 [0.42, 4.1] *	0.57 [-0.49, 1.63]
	Sasaki	3.07 [0.61, 5.53] *	1.30 [0.24, 2.36] *
	Hildebrand Hip	5.33 [3.91, 6.75] *	3.98 [3.30, 4.66] *
	Hildebrand Wrist	6.67 [4.53, 8.81] *	5.70 [3.53, 7.87] *
Recovery	Freedson	-3.10 [-3.91, -2.29] *	-4.41 [-5.24, -3.58] *
	Sasaki	-3.10 [-3.98, -2.22] *	-4.31 [-5.23, -3.39] *
	Hildebrand Hip	-3.75 [-4.48, -3.02] *	-4.34 [-5.22, -3.46] *
	Hildebrand Wrist	-4.37 [-5.15, -3.59] *	-5.14 [-5.96, -4.32] *
Kcals/interval	**Method**	**Protocol 1 Bias [95% C.I.]**	**Protocol 2 Bias [95% C.I.]**
Total	Freedson	-14.82 [-27.1, -2.54] *	-19.25 [-28.27, -10.23] *
	Sasaki	-11.77 [-26.55, 3.01]	-15.08 [-24.4, -5.76] *
	Hildebrand Hip	-8.15 [-15.98, -0.32] *	-1.83 [-8.8, 5.14]
	Hildebrand Wrist	-7.76 [-18.45, 2.93]	2.79 [-6.43, 12.01]
Work	Freedson	1.69 [0.32, 3.06] *	0.57 [-0.49, 1.63]
	Sasaki	2.30 [0.45, 4.15] *	1.30 [0.24, 2.36] *
	Hildebrand Hip	4.00 [2.94, 5.06] *	3.98 [3.30, 4.66] *
	Hildebrand Wrist	5.00 [3.40, 6.60] *	5.70 [3.53, 7.87] *
Recovery	Freedson	-4.66 [-5.88, -3.44] *	-4.41 [-5.24, -3.58] *
	Sasaki	-4.65 [-5.97, -3.33] *	-4.31 [-5.23, -3.39] *
	Hildebrand Hip	-5.63 [-6.72, -4.54] *	-4.34 [-5.22, -3.46] *
	Hildebrand Wrist	-6.55 [-7.72, -5.38] *	-5.14 [-5.96, -4.32] *

Table 2. kCal bias (ActiGraph EE estimate – criterion) by interval and method; *significantly different from zero ($p<0.05$).

In the current study, the underestimation of EE during recovery intervals was primarily responsible for the overall EE underestimation of the HIIT session for most methods. These findings may also apply to consumer wearables, which often produce EE underestimations (mean: -3%; median: -11%) when compared to indirect calorimetry.[6,7,27] The greatest underestimations produced by consumer wearables are reported during the highest intensity intermittent activities (e.g., basketball, tennis), suggesting that if worn for long periods at a time, highly active wearers would experience greater levels of EE underestimation than sedentary or low active wearers. If the results of the current study were extrapolated, a 60-minute interval-like activity of equivalent intensity could be underestimated by as much as 88 kCals (using the Freedson equation) and does not consider the slow phase of EPOC. Our results are directly applicable to the Freedson, Sasaki, and Hildebrand algorithms used with the ActiGraph but may apply to other research- and consumer-grade devices and algorithms.

This study had several strengths. Two separate protocols were completed in independent samples with similar results obtained from both. This suggests that regardless of the work-to-recovery ratio, accelerometer-based EE estimates from other HIIT protocols would reflect similar results to those in the current study. Additionally, the synchronization of the accelerometer and oxygen consumption data ensures that the current results are not due to discrepancies in data collection timing. Also, the wear locations make these results applicable to larger-scale epidemiology studies and to surveillance efforts using the ActiGraph.

One limitation of this study is that there was no long-term follow-up after the cessation of exercise to observe the slow component of EPOC on overall EE. Future studies could include a prolonged follow-up after interval training or other exercise or sports events. Also, both protocols had relatively small sample sizes (N=9 and N=20, respectively), which runs the risk that these results may not be replicated if the protocols were repeated with larger samples. Protocol 2 had an equal distribution of males and females, although protocol 1 only had one female participant, which could be considered a limitation. However, both protocols yielded similar results. The sample for both protocols included only healthy, young adults; therefore, it is possible that these results might not be applicable to other populations, such as sedentary adults, older adults, and children. Concerning the treadmill, safety was a priority, and so the protocols described in this study were limited to participant speeds at 95% of their HR max, which often translated to a fast jog or slow run for our participants. These results cannot be generalized to all types of HIIT training, such as body weight-centered circuit training or sprint interval training (defined as maximum all-out effort bouts).

CONCLUSIONS

In conclusion, this study affirms that estimations of EE converted from ActiGraph accelerometer data are underestimated during high-intensity interval training and identifies a specific source of error in the post-exercise time period that can be targeted in the future device calibration efforts. This error could be substantial for highly active individuals who perform interval-like exercise or sports, leading to discrepancies that could have meaningful over-or under-estimations of people's free-living behavior in many research studies that use the ActiGraph, and potentially other research- and consumer-grade activity trackers. Future research should 1) modify the intensity and duration of work and rest intervals to study the effects on EE estimates and 2) incorporate a longer post-exercise recovery period to more precisely estimate the contribution of the EPOC slow component to daily EE estimates. Continuing this line of research would further improve algorithms used to estimate EE from accelerometer data and could potentially be used to inform improvements to EE estimates of consumer activity tracking devices.

ACKNOWLEDGEMENTS

This research was funded by two University of Massachusetts Amherst Commonwealth Honors College Research Grants.

DECLARATION OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

1. Freedson PS, Melanson E, Sirard J. (1998) Calibration of the Computer Science and Applications, Inc. accelerometer. *Med Sci Sports Exerc*. 30(5):777-781. *doi:10.1097/00005768-199805000-00021*
2. Piercy KL, Troiano RP, Ballard RM, et al. (2018) The physical activity guidelines for Americans. *Jama*. 320(19):2020-2028.
3. Garriguet D, Tremblay S, Colley RC. (2015) Comparison of Physical Activity Adult Questionnaire results with accelerometer data. *Heal reports*. 26(7):11.
4. Lee IM, Shiroma EJ. (2014) Using accelerometers to measure physical activity in large-scale epidemiological studies: Issues and challenges. *Br J Sports Med*. 48(3):197-201. *doi:10.1136/bjsports-2013-093154*
5. Troiano RP, Berrigan D, Dodd KW, et al. (2008) Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc*. 40(1):181.
6. Adam Noah J, Spierer DK, Gu J, Bronner S. (2013) Comparison of steps and energy expenditure assessment in adults of Fitbit Tracker and Ultra to the Actical and indirect calorimetry. *J Med Eng & Technol*. 37(7):456-462.
7. Fuller D, Colwell E, Low J, et al. (2020) Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. *JMIR mHealth uHealth*. 8(9):e18694.
8. Panissa VLG, Fukuda DH, Caldeira RS, et al. (2018) Is oxygen uptake measurement enough to estimate energy expenditure during high-intensity intermittent exercise? Quantification of anaerobic contribution by different methods. *Front Physiol*. 9:868.
9. Knab AM, Shanely RA, Corbin KD, Jin F, Sha W, Nieman DC. (2011) A 45-minute vigorous exercise bout increases metabolic rate for 14 hours. *Med Sci Sport Exerc*. 43(9):1643-1648.
10. Batacan RB, Duncan MJ, Dalbo VJ, Tucker PS, Fenning AS. (2017) Effects of high-intensity interval training on cardiometabolic health: a systematic review and meta-analysis of intervention studies. *Br J Sports Med*. 51(6):494-503.
11. Norton K, Norton L, Sadgrove D. (2010) Position statement on physical activity and exercise intensity terminology. *J Sci Med Sport*. 13(5):496-502.
12. Buchheit M, Laursen PB. (2013) High-intensity interval training, solutions to the programming puzzle. *Sport Med*. 43(5):313-338.
13. Sawyer BJ, Tucker WJ, Bhammar DM, Ryder JR, Sweazea KL, Gaesser GA. (2016) Effects of high-intensity interval

training and moderate-intensity continuous training on endothelial function and cardiometabolic risk markers in obese adults. *J Appl Physiol.* 121(1):279-288.
14. Wewege M, Van Den Berg R, Ward RE, Keech A. (2017) The effects of high-intensity interval training vs. moderate-intensity continuous training on body composition in overweight and obese adults: a systematic review and meta-analysis. *Obes Rev.* 18(6):635-646.
15. Zhang H, Tong TK, Qiu W, et al. (2017) Comparable effects of high-intensity interval training and prolonged continuous exercise training on abdominal visceral fat reduction in obese young women. *J Diabetes Res.* 2017.
16. Keating SE, Johnson NA, Mielke GI, Coombes JS. (2017) A systematic review and meta-analysis of interval training versus moderate-intensity continuous training on body adiposity. *Obes Rev.* 18(8):943-964.
17. Godin G. (2011) The Godin-Shephard leisure-time physical activity questionnaire. *Heal & Fit J Canada.* 4(1):18-22.
18. Warburton DER, Gledhill N, Jamnik VK, et al. (2011) Evidence-based risk assessment and recommendations for physical activity clearance: Consensus Document 2011. *Appl Physiol Nutr Metab.* 36(S1):S266--S298.
19. Åstrand P-O, Åstrand I. (1958) Heart rate during muscular work in man exposed to prolonged hypoxia. *J Appl Physiol.* 13(1):75-80.
20. Powers SK, Howley ET, Quindry J. (2007) *Exercise Physiology: Theory and Application to Fitness and Performance.* McGraw-Hill New York, NY.
21. Weir JB de V. (1949) New methods for calculating metabolic rate with special reference to protein metabolism. *J Physiol.* 109(1-2):1-9.
22. Sasaki JE, John D, Freedson PS. (2011) Validation and comparison of ActiGraph activity monitors. *J Sci Med Sport.* 14(5):411-416. *doi:10.1016/j.jsams.2011.04.003*
23. Hildebrand M, Van Hees VT, Hansen BH, Ekelund U. (2014) Age group comparability of raw accelerometer output from wrist-and hip-worn monitors. *Med Sci Sports Exerc.* 46(9):1816-1824. *doi:10.1249/MSS.0000000000000289*
24. Crouter SE, Churilla JR, Bassett DR. (2006) Estimating energy expenditure using accelerometers. *Eur J Appl Physiol.* 98(6):601-612. *doi:10.1007/s00421-006-0307-5*
25. Imboden MT, Nelson MB, Kaminsky LA, Montoye AH. (2018) Comparison of four Fitbit and Jawbone activity monitors with a research-grade ActiGraph accelerometer for estimating physical activity and energy expenditure. *Br J Sports Med.* 52(13):844-850. *doi:10.1136/bjsports-2016-096990*
26. Lyden K, Kozey SL, Staudenmeyer JW, Freedson PS. (2011) A comprehensive evaluation of commonly used accelerometer energy expenditure and MET prediction equations. *Eur J Appl Physiol.* 111(2):187-201.
27. Sasaki JE, Hickey A, Mavilia M, et al. (2015) Validation of the Fitbit wireless activity tracker for prediction of energy expenditure. *J Phys Act Heal.* 12(2):149-154.

ABOUT THE STUDENT AUTHORS

This data was collected, and the majority of the manuscript was written by Nicholas Remillard and Marisa Mulvey during their senior years of their undergraduate studies. As his undergraduate honors thesis, Nicholas Remillard developed the research question and study protocols under the guidance of his mentor John Sirard and master's student Greg Petrucci. Marissa Mulvey continued data collection and improved upon the original study protocol for her undergraduate honors thesis, and both Nicholas Remillard and Marissa Mulvey collaborated on data analyses and the writing of the manuscript.

PRESS SUMMARY

Measuring energy expenditure is critical in understanding physical activity's relationship with health and disease. Currently, there is no way to directly measure energy expenditure outside of laboratory settings. Existing energy expenditure prediction algorithms based on wearable devices inaccurately estimate energy expenditure outside of controlled laboratory conditions. The current study identified the post-exercise time period as a source of underestimation of energy expenditure from the wearable device algorithms. This underestimation could be substantial for highly active individuals who perform interval-like exercise or sports and rely on a wearable device to track their calories burned. In addition, the results of this study demonstrate current limitations that should be considered by researchers studying the metabolic effects of high-intensity exercise using wearable devices.