2014

Relative validity of 3 accelerometer models for estimating energy expenditure during light activity

Alexander Allan Wetten
University of Wollongong, aaw153@uowmail.edu.au

Marijka Batterham
University of Wollongong, marijka@uow.edu.au

Sze Yen Tan
University of Wollongong, syt153@uow.edu.au

Linda Tapsell
University of Wollongong, ltapsell@uow.edu.au

Relative validity of 3 accelerometer models for estimating energy expenditure during light activity

Abstract

Background: With physical inactivity inextricably linked to the increasing prevalence of obesity, there is a need for validated methods that measure free-living energy expenditure (EE) within sedentary environments. While accelerometers enable these measurements, few studies have compared device accuracy in such settings. The aim of this study was to investigate the relative validity of the Actigraph, RT3 and SenseWear Armband (SWA). Methods: Twenty-three (11 male, 12 female) participants (age: 25.3 ± 6.3 yr; BMI: 22.6 ± 2.7) wore 3 accelerometers at designated sites during a 4-hour stay in the Whole Room Calorimeter (WRC). Participants performed 2 10-minute bouts of light-intensity exercise (stepping and stationary cycling) and engaged in unstructured sedentary activities. EE estimated by accelerometers was compared with WRC EE derived from measurements of gaseous exchange. Results: The Actigraph and SWA both accurately estimated EE during the stepping exercise. EE estimated by the RT3 during stepping was significantly lower than the WRC value (31.2% ± 15.6%, $P < .001$). All accelerometers underestimated cycling and unstructured activity EE over the trial period ($P < .001$). Conclusions: The Actigraph and SWA are both valid tools for quantifying EE during light-intensity stepping. These results provide further valuable information on how accelerometer devices may be appropriately used

Keywords

estimating, models, accelerometer, validity, relative, activity, light, during, expenditure, energy, 3

Disciplines

Education | Social and Behavioral Sciences

Publication Details

Wetten, A. Allan., Batterham, M., Tan, S. & Tapsell, L. (2014). Relative validity of 3 accelerometer models for estimating energy expenditure during light activity. Journal of Physical Activity and Health, 11 (3), 638-647.

Relative validity of 3 accelerometer models for estimating energy expenditure during light activity

Alexander Allan Wetten, Marijka Batterham, Sze Yen Tan, and Linda Tapsell

Affiliations: Wetten, Tan and Tapsell are with School of Health Sciences, Smart Foods Centre, University of Wollongong, Wollongong, NSW, Australia. Batterham is with Statistical Consulting Service, University of Wollongong, Wollongong, NSW, Australia.

Running title: Validity of three accelerometer models

Manuscript type: original research

Key Words: accelerometry, whole room calorimeter, sedentary, accuracy.

Abstract word count: 199 words

Manuscript word count: 5757

Date of submission: 10th September, 2012

ABSTRACT

Background With physical inactivity inextricably linked to the increasing prevalence of obesity, there is a need for validated methods that measure free-living energy expenditure (EE) within sedentary environments. While accelerometers enable these measurements, few studies have compared device accuracy in such settings. Our aim was to investigate the relative validity of the Actigraph, RT3 and SenseWear Armband (SWA).

Methods Twenty-three [11 male, 12 female] participants [age: 25.3 ± 6.3 yr; BMI: 22.6 ± 2.7] wore three accelerometers at designated sites during a four-hour stay in the Whole Room Calorimeter (WRC). Participants performed two 10-minute bouts of light-intensity exercise (stepping and stationary cycling) and engaged in unstructured sedentary activities. EE estimated by accelerometers was compared with WRC EE derived from measurements of gaseous exchange.

Results The Actigraph and SWA both accurately estimated EE during the stepping exercise. EE estimated by the RT3 during stepping was significantly lower than the WRC value (31.2% ±15.6%, $P < 0.001$). All accelerometers underestimated cycling and unstructured activity EE over the trial period ($P < 0.001$).

Conclusions The Actigraph and SWA are both valid tools for quantifying EE during light-intensity stepping. These results provide further valuable information on how accelerometer devices may be appropriately used.

INTRODUCTION

Paragraph 1: Obtaining an accurate assessment of physical activity is an essential component of obesity research. A growing body of research has shown that many individuals worldwide are physically inactive (1), which is inextricably linked to the increasing prevalence of global overweight and obesity and associated with an increased risk of chronic disease (2, 3). Physical activity (PA) comprises both planned, structured exercise and spontaneous incidental activity built up incrementally over the day (4). With growing recognition of the health benefits attained through reducing physical inactivity (screen time, sitting etc.) and increasing incidental physical activity (e.g. taking the stairs instead of the elevator or cycling to work rather than driving)(5), there is an increasing need in research to be able to quantify modifiable aspects of energy expenditure (EE) from lighter intensity PA and sedentary behavior using valid methods (6).

Paragraph 2: Accelerometers have become increasingly utilized in obesity research and practice (7-9). Their advantages include convenience, small size and ease of wear, they are relatively inexpensive, provide real-time data acquisition, and can be used in both the research setting and free-living environment (8, 10, 11). Prior research has shown that accurate estimation of EE from accelerometer output (usually activity counts) depends on many factors, including: the type of regression equation used (12), amount of time for which the device is worn (7), number of axes employed (13), placement site selected (9), population demographics and body composition (8) and physical activity type (14). A major limitation of accelerometry however is that not all types of physical activity are accurately characterized by the activity counts obtained by most devices. Many accelerometers have difficulties discerning upper body and arm movements, which include weight-bearing activities, stepping up on an incline and cycling (15). Recent accelerometer innovations have attempted to integrate sensory information in an attempt to overcome these limitations (16), but further recognition is needed in research.

Paragraph 3: While several studies have validated the use of accelerometers to predict energy expenditure in moderate to high-intensity PA exercise conditions, there is a paucity of relative validity studies where accelerometers are compared to a reference method in the research literature for lighter intensities and sedentary activity. The few studies which have assessed the relative validity of several different accelerometers under sedentary and light PA conditions have shown conflicting results (11). While some studies have shown accelerometers to be accurate for predicting EE when appropriate regression equations were utilized (12, 13), others have reported over- (17, 18) and underestimations of EE (19-21). Thus, the ability of accelerometers to accurately quantify the energy cost of sedentary and lighter-intensity PA remains in question (11, 22). There is a need for further validation studies to assess the role of accelerometers in the context of sedentary and lighter PA, which reflects the lifestyle of overweight and obese individuals.

Paragraph 4: The aim of this study was to investigate the relative validity of three different commercially available accelerometers (Actigraph GT1M, RT3 and SenseWear Armband) for estimating energy expenditure in a sedentary environment, compared simultaneously with whole room calorimetry as a reference method. Furthermore, accelerometers used in this study each had a different number of measurement axes and were placed at different body locations.

METHODS

Paragraph 5*:* This was a validation study comparing EE predicted by three different accelerometers with EE measured in a whole room calorimeter (WRC) facility over a four-hour period in normal healthy adults. Participants were required to wear all three accelerometers during the calorimeter stay. Measured (WRC) and predicted (accelerometer) energy expenditure were compared to establish the accuracy and relative validity compared to the WRC of each accelerometer, throughout the trial period and while participants performed two 10-minute low-intensity exercises (stepping and cycling). The protocol used in this study was approved by the University of Wollongong Ethics Committee (HE09/208).

Participants

Paragraph 6: Adult participants were recruited from the general population of staff and students at Wollongong University via flyers and announcements. All participants reported being free of metabolic illnesses and chronic conditions (e.g. diabetes and thyroid disorders) and were not taking any medications known to affect energy metabolism. Informed written consent was obtained from all participants after an explanation of the purpose and procedures of the study.

Experimental protocol

Paragraph 7: Prior to the calorimeter stay, participants attended an information session where habitual physical activity (23) and dietary intake were assessed (24). Height, weight and percentage of body fat were measured (in light clothing, without shoes) using a stadiometer and leg-to-leg bioelectrical impedance scales (Tanita Corp, model UM019, Tokyo). Body mass index was calculated and waist and hip measurements were taken to assess regional fat distribution. Participants were then shown the WRC facility, to ensure they were comfortable with the stay in the chamber.

Paragraph 8: For the calorimeter visit, participants were asked to fast and restrict caffeine for at least 10 hours, as well as refrain from strenuous physical activity the day before. All participants were inside the WRC for four hours; entering between 0800 and 0830 and exiting from 1200 to 1230. Prior to entering the WRC, each participant had all three accelerometers attached at designated hip and arm sites. The Actigraph GT1M (ACT) was secured randomly to the right or left hip at the anterior superior iliac spine, with the RT3 positioned contralateral to the ACT. The SenseWear Armband (SWA) was positioned on the tricep of the dominant arm. Each device was pre-programmed with information about gender, age, height and weight of participants, to allow for EE estimation. One hour after entering the WRC, participants were provided with breakfast based on the participants' usual intake and energy needs, calculated to meet 30% of daily energy requirements, using predictive equations (25). The macronutrient profile was standardized (15% protein, 30% fat, 55% carbohydrate) for each participant using FoodWorksv7.0 software (Xyris Software, Brisbane, Australia, Professional Edition).

Paragraph 9*:* While inside the WRC, participants were required to perform two 10-minute exercises (stepping and stationary cycling) at a low intensity, which provided structured activities to allow for a comparison between accelerometers. These activities were selected, as both have previously been validated with an accelerometer at greater intensities (12, 26) and options for different exercise protocols within the WRC are limited due to the constrained environment. Exercises were performed at 120 minutes (stepping) and 170 minutes (cycling). The *stepping exercise* involved stepping up (step height 233mm) to a metronome set at 40 beats per minute and stepping down to the subsequent beat, equating to 20 steps per minute. Instructions were provided to allow the arms to swing naturally. The *cycling exercise* involved stationary cycling for 10-minutes at 60rpm, with a preset intensity of 50 watts (Monark, Ergomedic 828E, Sweden). Outside the structured exercise protocol, participants were free to engage in any sedentary activities. Sedentary activities available were: sitting in a chair watching television, using a computer, or desk work, which involved writing or reading. Participants were free to mobilise within the WRC, though were discouraged from engaging in any physically demanding activity. Following the completion of each calorimeter stay, height, weight and body fat were re-recorded.

Instrumentation

Paragraph 10*:* The three accelerometers used in this study were the Actigraph GT1M [ACT], RT3 and SenseWear Armband [SWA]. The ACT (Actilife v.4.1.1 Firmware v.3.2.0, Pensacola, FL, USA) is a dual-axial accelerometer (38mm x 37mm x 18mm; 27g) measuring accelerations in the vertical and horizontal planes by means of a solid-state monolithic sensor (27). ACT output is digitized by a twelve-bit Analog to Digital Convertor (ADC) at a rate of thirty times per second (30 Hertz) and detects acceleration ranging from 0.05 to 2.5 G. The acceleration signal is filtered (0.25–2.5Hz), rectified and integrated in a capacitor. The RT3 accelerometer (Stayhealthy Inc., v1.2, Assist v1.0.7, Monrovia, CA, USA) is a tri-axial accelerometer (71mm x 56mm x 28mm; 65.2g, including battery) measuring vectors in the vertical, anteroposterior and mediolateral planes. The SWA (SenseWear Professional 3, Bodymedia, Inc., v6.1.0, Pittsburgh, PA, USA) is a wireless, multisensory activity monitor worn over the triceps muscle on the dominant arm. The SWA integrates data from a dual axial accelerometer, galvanic skin response sensor (GSR), heat flux sensor, skin temperature sensor and near-body ambient temperature sensor, to estimate EE under free-living conditions. The SWA continually updates its software to calculate activity specific algorithms, though the direct contribution of each sensor to predict EE is not published (16).

Whole room calorimeter (WRC)

Paragraph 11*:* The calorimeter facility located at University of Wollongong measures oxygen consumption and carbon dioxide production through airtight, ventilated and air-conditioned chambers. Details of the protocols and operating conditions have been previously published (28). Rates of oxygen consumption (Vo_2) and carbon dioxide production (Vco_2) were calculated by the flow rate of gases out the chamber and the concentrations of inlet and outlet air from the chamber, according to Schoffelen et al (29). EE was calculated through gaseous exchange using the Weir equation (30). Prior to each visit, the gas analyzers were calibrated and the accuracy and precision regularly tested by alcohol combustion.

Data processing and statistical procedures

Paragraph 12*:* Both the RT3 and SWA utilized proprietary manufacturer equations to predict EE, while the Actigraph offers the Work-Energy Theorem and Freedson Equation (27, 31) to estimate EE. For this study, we applied Crouter's (12) 2-way regression equation for Actigraph values for the following reasons: *i)* it has previously been shown to provide greater accuracy under the state of light-intensity exercise (12, 19); *ii)* more recent studies have shown that linear regression models poorly predict energy expenditure from accelerometer output (32, 33). While a revised Crouter equation exists to prevent misclassification of activities commencing in the middle of an ACT minute (34), a recent study has shown that this equation underestimates EE for lighter activity when compared with the initial equation (35). This was also apparent during preliminary testing in our study, so the revised equation was not used. Activity counts in the vertical axis for the Actigraph were stored in 10s epochs to allow for Crouter's EE equation, and subsequently transformed into one-minute epochs for a comparison between accelerometers. All data for the RT3 and SWA were collected in one-minute epochs. The WRC samples room air every two minutes, and averages data into 10-minute epochs to calculate EE. The initial and final 10-minutes of WRC data were not used in the analyses, to allow for artifact from participants entering and exiting the chamber.

Paragraph 13*:* Relative validity was assessed by comparing absolute EE to the WRC values for both exercises and throughout the trial period, as well as on a minute-by-minute basis, using the following procedures:

i) Totals were calculated for participant EE during 10-minutes of stepping and cycling exercise, and for unstructured activity with the remainder of the trial (200 minutes of data averaged for analysis into a 10 minute period for comparison with the other activities).

ii) To determine the relative validity of each device compared with WRC, Bland-Altman plots of the difference in EE estimated by each accelerometer and EE measured by WRC were used (36). Percentage difference was calculated as: [(predicted EE – measured EE)/measured EE] x 100.

iii) Data were further analyzed for minute-to-minute differences in EE between each accelerometer for both exercises. However, as the EE data from the WRC is calculated for each 10-minute time period, a comparison to the reference could not be made for the exercises.

Paragraph 14*:* In order to compare the periods of stepping, cycling and unstructured activity in a single analysis, a non-exchangeable multivariate hierarchical Bayesian model was used (37). Classical (or frequentist) methods of statistical analysis assume that each individual study is one in a long running series of experiments in which the current study estimates are likely to lie within the stated confidence intervals 95% of the time, these methods assume that only repeatable experiments have a probability. Bayesian methods ascribe a probability distribution to the study estimate which reflects our prior belief and about the mean combined with the study data, in this framework the probability reflects uncertainty, both the uncertainty related to random sampling as recognized by the classical (frequentist) framework and the uncertainty of not knowing the true value. In the Bayesian framework the unknown population parameter is modeled with a probability distribution rather than being considered as a fixed (unknown) single value as in the classical framework (38). Bayesian methods have advantages over frequentist methods in the natural incorporation of hierarchical data, missing data and their effectiveness with estimation in small sample sizes all of which provide advantages in physical activity research (37, 39, 40). In this hierarchical model, there are repeated measurements by the four methods made on the same subjects, where there is an assumption that the underlying value of the measurement could be continually changing. In this case, the estimate of limits of agreement is made by modeling the paired differences (41). Units employed are the 10-minute period of stepping, the 10-minute period of cycling and the remaining period of 200 minutes averaged over a 10-minute period. As measurements in this type of analysis can be highly correlated we used a large number of simulations, multiple chains, over-relaxation and a substantial burn in period to reduce any effect of autocorrelation. One hundred thousand simulations were run in four parallel chains (with over-relaxation) with the first 5,000 iterations discarded as a burn in period allowing stabilization of the model (WinBUGS version 1.4 MRC Cambridge (42)). Models were checked for convergence using the trace history plots of the simulations. The median and 95% credible region are reported, which is equivalent to the 2.5 and 97.5 percentiles of the posterior distribution. A repeated-measures MANOVA was used to assess difference between methods (ACT, RT3, SWA and WRC) and the three activity periods (stepping, cycling and unstructed activity). The repeated measures MANOVA accounts for the correlation between the energy expenditure assessed using the four different methods.

This model is used to compare the individual activities as distinct from the hierarchical model which compares the whole trial period. Analyses were conducted using SPSS for Windows (V17.0, SPSS, Chicago, IL) and two-tailed statistical significance set at $p<0.05$.

RESULTS

Participant Demographics

***Paragraph 15**:* Twenty-three (12 female, 11 male) healthy, adult participants completed the study. Participant physical characteristics are displayed in Table 1. Software error resulted in Actigraph data loss for one participant.

Table 1. *Physical characteristics of participants (n=22)*

Variable	Mean ± SD (total range)
Age (years)	25.3 ± 6.3 (19–43)
Height (m)	1.75 ± 0.84 (1.59–1.90)
Weight (kg)	69.3 ± 10.1 (50.9–94.5)
BMI ($kg.m^{-2}$)	22.6 ± 2.7 (17.1–28.0)
Body Fat (%)	21.8 ± 7.3 (10.7–36.0)
Waist (cm)	75.4 ± 7.7 (63.3–92.0)
W:H ratio	0.80 ± 0.06 (0.66–0.89)

BMI, body-mass index; W:H, waist to hip.

Absolute differences in EE

***Paragraph 16**:* To compare each accelerometer with the WRC over the whole (220 minute) trial, a multivariate non-exchangeable Bayesian analysis was conducted. This allowed for all information to be analyzed simultaneously (with the addition of the missing ACT data from the one subject), while providing greater power and increased statistical robustness. In this framework, the WRC was significantly different to EE predicted by all accelerometers (Table 2). The SWA and ACT accelerometers were not significantly different to one another, while the RT3 gave lower EE results than both the other accelerometers.

Table 2. *Estimates of bias and the median (kcal/10min) with associated 95% credible regions (95% CR) for pair-wise comparisons of EE measured by the WRC and three accelerometers (ACT, RT3 and SWA) over the whole trial.*

Pair-wise comparisons	Bias (kcal/10min)[a]	Median (95% CR), kcal/10min
ACT – WRC	-4.23 ± 1.30	-4.24 (-6.79, -1.68)*

RT3 – WRC	-9.96 ± 1.00	-9.97 (-11.93, -8.01)*
SWA – WRC	-5.80 ± 1.13	-5.78 (-8.10, -3.65)*
RT3 – ACT	-5.73 ± 1.22	-5.73 (-8.15, -3.38)*
SWA – ACT	-1.56 ± 1.17	-1.55 (-3.86, 0.72)
SWA – RT3	4.16 ± 1.06	4.16 (2.09, 6.27)*

ACT, Actigraph; SWA, SenseWear Armband; WRC, Whole Room Calorimeter.

[a]Values are means ± SD. *Results which do not contain 0 in the CR are significantly different (α<0.05) using a multivariate hierarchical Bayesian model.

Table 3. *Predicted EE for the ACT, RT3 and SWA accelerometers compared to EE measured by the Whole Room Calorimeter*

	Activity		
Method	Stepping Exercise (kcal/10min)	Cycling Exercise (kcal/10min)	Unstructured Activity (kcal/200min)
ACT	40.11±6.83	22.88±14.25*	287.23±52.24*
RT3	27.40±6.35*	19.94±4.71*	262.21±36.23*
SWA	38.94±7.66	19.92±6.18*	280.12±38.53*
WRC	40.15±8.03	33.99±6.87	338.75±58.87

Values are means (SD).

ACT, Actigraph, SWA, SenseWear armband, WRC, Whole room calorimeter, Unstructured Activity, Whole trial minus stepping and cycling exercises.

*Significantly different from WRC, $P < 0.05$.

Table 4. *Percentage difference of energy expenditure estimated by ACT, RT3 and SWA compared with EE measured via the Whole Room Calorimeter*

	Percentage Difference[a]		
Method	Stepping Exercise	Cycling Exercise	Unstructured Activity
ACT - WRC	0.1	-35.1	-14.7

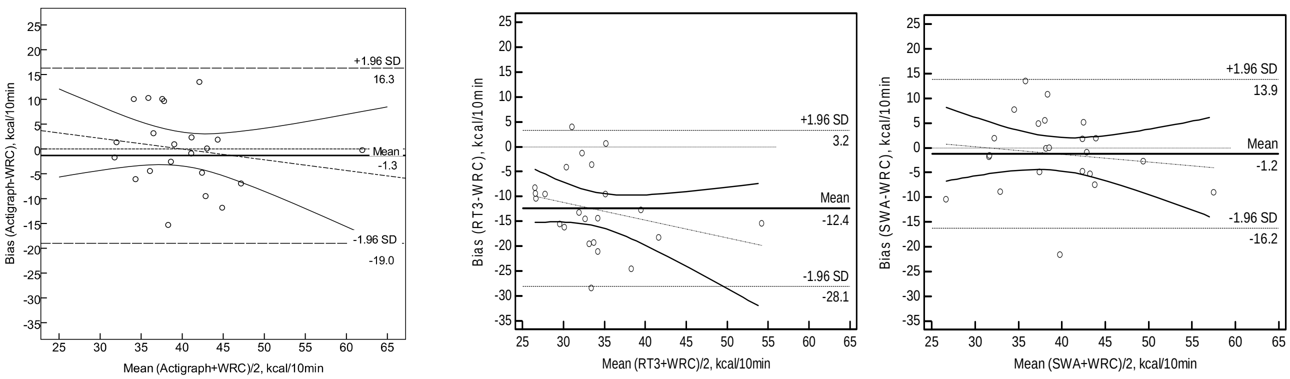
RT3 - WRC	-31.7	-62.7	-22.7
SWA - WRC	-3.1	-68.9	-17.6

ACT, Actigraph, SWA, SenseWear Armband, WRC, Whole Room Calorimeter.

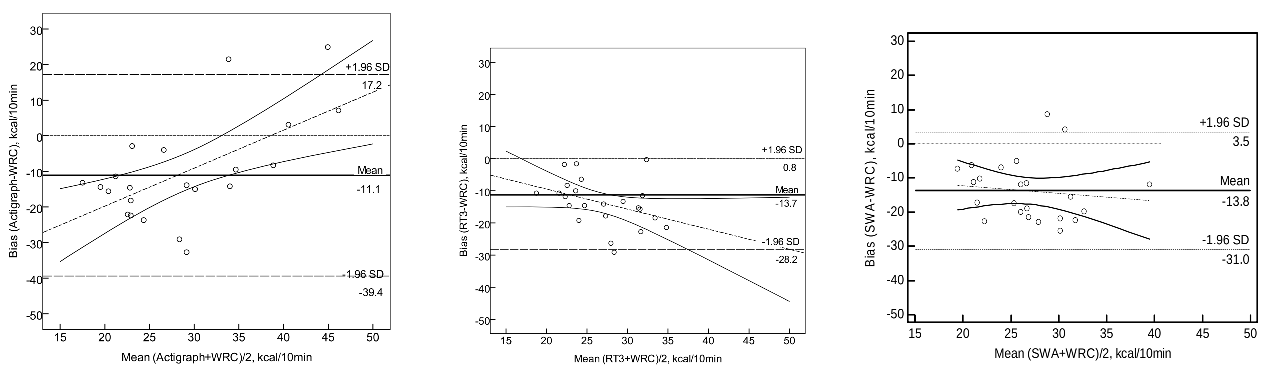
[a]Percentage difference: [(predicted EE – measured EE)/measured EE] x 100.

Paragraph 17: EE measured by the WRC and predicted by accelerometers (in kcal) are presented in Table 3, differences expressed as a percentage are presented in Table 4. The repeated-measures MANOVA showed estimates of EE between accelerometers and the WRC during stepping demonstrated no significant difference for the ACT or SWA when compared with the WRC. RT3 stepping EE was 31.2% lower than the WRC, $P<0.05$). For cycling and unstructured activity, all 3 accelerometers significantly underestimated EE ($P < 0.05$). The correlation coefficients between the WRC and the ACT, RT3, and SWA are as follows respectively: stepping, 0.488, 0.400 and 0.551; cycling, 0.177, 0.209, and 0.071; sedentary activity, 0.638, 0.725, and 0.800; and for the overall study 0.613, 0.680, and 0.809.

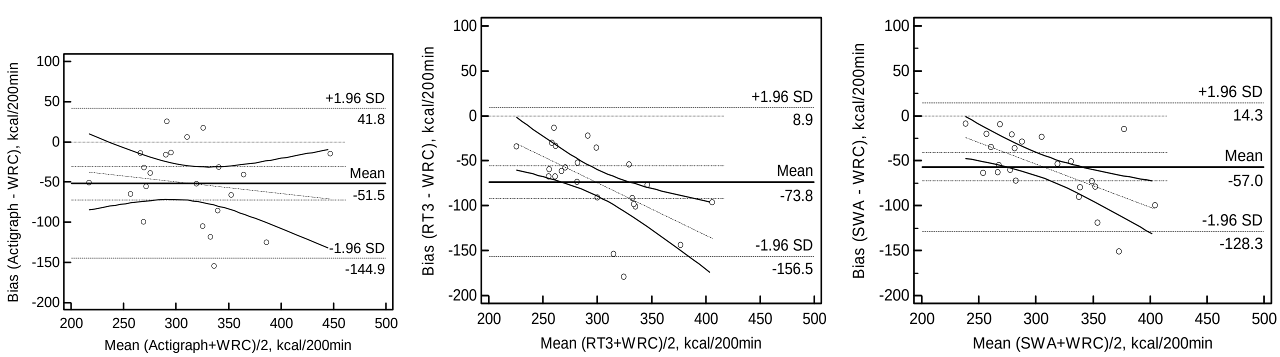
Paragraph 18*:* Distributions of bias for each accelerometer compared to the WRC are presented as Bland-Altman plots (Figure 1). For the stepping exercise (Figure 1A), the small bias, indicated by scores clustered close to the zero mark, and low percentage difference for the ACT and SWA illustrate a high degree of accuracy for estimating EE with this exercise. In contrast, the large bias and majority of scores bellow the zero line on the *y*-axis for the RT3 shows that this device tended to consistently underestimate for light-intensity stepping. For the cycling exercise and the unstructured activity, all devices exhibited a large degree of bias and a loose distribution of scores with wide limits of agreement, indicating poor precision and substantial underestimation of EE (on average > 11kcal/10mins for cycling; >51kcal/200mins for the unstructured activities).



A - Stepping



B - Cycling



C- Unstructured Activity

Minute-to-minute EE comparisons for exercises

Paragraph 19*:* A comparison of minute-to-minute EE estimated by the ACT, RT3 and SWA during the exercises revealed a significant difference between devices for stepping ($F_{2,65}$ = 17.31, P < 0.001), with an interaction between method and time also observed ($F_{9,11}$ = 2.07, P = 0.01). *Post hoc* Bonferroni testing revealed that EE estimated by the RT3 was significantly lower than the ACT (95% CI = 0.59 – 1.70, P <

0.001) and SWA (95% CI = 0.58 – 1.67, *P* < 0.001), (mean difference = 1.15 ± 0.22 and 1.13 ± 0.22 respectively).

DISCUSSION

Paragraph 20: This study produced varied results depending on the type of activity. The ACT and SWA devices produced more accurate readings compared to the RT3, and were shown to be valid gauges for estimating EE during light-intensity stepping. All accelerometers exhibited poor capacity to predict EE from stationary cycling, and each device significantly underestimated EE from sedentary, unstructured activities over the remaining trial time period. On a minute-to-minute basis, the RT3 significantly underestimated stepping EE compared to the ACT and SWA devices.

Stepping exercise

Paragraph 21: The closest agreement between accelerometers and the WRC values was observed for stepping, with values from both the ACT and SWA devices within 3.1% of the measured EE value from the WRC. A previous study reported comparable accuracy with uniaxial accelerometry during a bench stepping exercise, using revised regression models (12). Other studies with have found accelerometers to underestimate EE during stepping (43) and stair walking exercises (19), though this difference may have been be due to our slower stepping protocol (20 steps/min), where we found the best performing device was comparable to the WRC (0.1% difference). Trends in underestimating the energy cost of stepping and stair walking seen in other studies (18, 44, 45) may be partly attributed to greater workloads and the additional EE required to ascend a staircase, as opposed to bench stepping. It has also been suggested that underestimation during stepping may result from an inability for accelerometers to account for vertical displacement, with the EE required to generate force during incline stepping being larger than the displacement measured (44). Another recent study confirmed that underestimation is likely occurring in the ascending phase (which requires greater effort) and overestimation ensuing during the subsequent decent (46).

Paragraph 22: Although prior research has recommended that accelerometers should be applied for differentiating activity levels rather than providing absolute estimates of EE (45), the high degree of accuracy observed for both the ACT and SWA in this study suggests that these devices can be used to estimate EE from stepping. However, while a high level of accuracy was observed within the study group, a large range did exist between individuals. In all, our results show that both hip and arm placed accelerometers are useful gauges for predicting EE during the light-intensity stepping exercise.

Cycling exercise

Paragraph 23: Our finding that the ACT, RT3 and SWA all exhibited poor predictive ability during the light-intensity cycling exercise is in agreement with several other trials (16, 44, 45). We found one study that did find the SWA provided valid estimates of EE during stationary cycling exercise, when subjects pedaled consistently at 60rpm at 60% of their predetermined Vo_{2peak} (26). In this study, subjects burnt an average of 93.0kcal per 10 minutes vs 34.0kcal used in our trial, suggesting that higher PA intensity may have led to greater accuracy for this exercise. Overall however, underestimation of EE has most frequently been observed using a range of devices during stationary cycling, despite the employment of alternative algorithms and variations in device placement. These collective results suggest that at present, accelerometers are not appropriate tools to accurately quantify EE during light-intensity stationary cycling.

Whole trial EE

Paragraph 24: In agreement with most studies, our study found that accelerometers are poor at estimating the EE of sedentary lifestyle activities, particularly when upper body movements predominate (20, 46). Most accelerometers are unable to detect certain postural changes that would likely yield physiological differences in energy expenditure (47). As there were considerable arm movements during the trial e.g. using a keyboard and eating breakfast, we anticipated that the SWA armband placed on the upper body would detect this to a greater extent than the two hip-placed devices, providing more accurate estimates of EE. It was further predicted that the additional SWA sensors, which detect temperature and galvanic skin response changes, would improve the overall validity of the device. However, our results demonstrated that incorporating sensory data with an upper body accelerometer did not lend further accuracy for estimating EE in a sedentary, unstructured environment.

Paragraph 25: A few studies have shown uniaxial accelerometers to be valid devices for estimating EE in sedentary environments. One study investigated the validity of an earlier Actigraph model (7164) to predict EE from a number of sedentary activities, such as filing papers and computer work (12). Using a novel algorithm, these devices were shown to be accurate in this setting. Another recent study demonstrated that uniaxial accelerometers can accurately predict EE from spontaneous low-intensity habitual activities from participants within a respiratory chamber (48). In our study, there was a significant difference between the measured WRC values and the (uniaxial) ACT accelerometer, but this may reflect our young, healthy study sample and the exercises performed. In all, the average difference between our measured (WRC) and predicted (ACT) value may have only minor clinical relevance in shorter, population-based studies that use the ACT to estimate absolute EE from sedentary activity.

Paragraph 26: Prior research has proposed that additional axes of measurement during lifestyle activities may be beneficial, due to an increased ability to detect horizontal movements which predominate in sedentary settings (18). However, in line with our findings, a number of studies investigating this assumption have reported little improvements in accuracy using three planes of reference compared to uniaxial accelerometry (20, 46, 49). While utilizing additional axes did not improve EE estimations for the accelerometers in our study, one recent study did find that accuracy was improved using the AC-210 triaxial accelerometer. (13). In this study, 21 Japanese male adults performed low-level and sedentary PAs during a 22.5 hour WRC stay. The authors reported that these devices provided accurate estimates when the horizontal plane was used to differentiate lifestyle activities, in combination with novel prediction equations derived from participants sleeping metabolic rate. Further investigation is required to ascertain the extent to which improvements in accuracy can be attributed to the device used, equations employed and the number of planes of reference.

Paragraph 27: Another potential explanation for the poor estimation of sedentary activities may be due in part to differences in sensor technology, sampling frequency and bandpass filtering (50), as some accelerometers have difficulty registering slower ambulatory movements (11). In comparisons between three Actigraph accelerometer models using mechanical oscillations, the recent GT1M Actigraph was found to be poorly suited for lighter activity, due to lower monitor sensitivity and a higher threshold for non-zero counts, compared to earlier models (22). In our study, it is possible that the frequency response of the Actigraph (0.25 – 2.5 Hz) was unable to detect sedentary behaviors accurately. Further research is needed to determine how monitor sensitivity and filtering affects prediction of EE for lighter activity.

***Paragraph 28**:* The ability to accurately predict EE from exercise and incidental activity is dependant on the regression algorithm used with the accelerometer (50). Researchers utilising novel prediction equations to characterize different types of physical activity have consistently reported greater accuracy than when proprietary equations are used (12, 13). In our study, the ACT was the only accelerometer whereby the prediction equation could be user specified with the software. When the Freedson and Work-Energy Theorem equations were used, the Actigraph significantly underestimated EE for each exercise condition (data not shown), confirming findings in earlier studies (12, 17). Crouter's equation (12), which uses differences in the coefficient of variation on a 10 second interval basis to distinguish locomotor and non-locomotor activity, significantly improved the accuracy of estimations in the current study. Preliminary testing with Crouter's refined equation (34) however showed EE was significantly underestimated for both exercises and unstructured activity. Further work is needed to develop and refine equations that accurately characterize the type and energy cost of sedentary activity.

Paragraph 29: The limitations of our study are reflected in our study sample and the possible presence of inter-monitor variability in measurements. The devices have been used in a number of trials, but had not been recently calibrated. The determination of calibration offset factors for the SWA and RT3 would reduce this small source of error. The study sample comprised young, fit and healthy adults, and this limits the generalizability of results to others of different body mass, fitness level or age. Some factors that can influence EE, such as daily exercise and menstrual cycle (8), were not controlled for, however with the length of the measurement period, this was not likely to bias results. In future, the use of larger study samples would allow for multiple regression modeling to determine whether fat mass and fat-free mass impart a bias on EE predicted by accelerometers. The development of accurate equations and site selection to allow prediction of EE from cycling is another future research avenue.

CONCLUSIONS

Paragraph 30: This study was one of only a few studies investigating the relative validity of commercially available accelerometers with different numbers of measurement axes in a sedentary environment. It has shown that for the accelerometers used in this trial, incorporating additional sensory data and planes of reference do not lend further accuracy to estimates of EE during light exercise and unstructured activity. While the ACT and the SWA devices produced the most valid estimates of EE during the stepping exercise, all devices were poor at estimating EE from stationary cycling and unstructured, sedentary activity within the WRC. The high variability in individual scores evidenced during Bland-Altman testing suggests that accelerometers are best utilized for the assessment of stepping at a population, rather than individual level. Overall, the results from this study provide a valuable contribution to the small body of literature examining the accuracy of different accelerometers within controlled, sedentary settings. Additional validation studies are warranted in natural, real life environments to determine the appropriate use of accelerometers with inactive populations and provide further understanding of the relationships between obesity, physical activity and health.

Acknowledgements

We thank all participants who volunteered time and energy for this study. We would also like to thank Dr Tanya Strevens for editing the manuscript.

Funding source

This study was funded by the Smart Foods Centre, University of Wollongong. Results from the current study do not constitute endorsement by the International Society for Physical Activity and Health. There is no conflict of interest and the authors declare no professional relationships with any companies or manufacturers who may benefit from the studies results.

REFERENCES

1. Hansen BH, Kolle E, Dyrstad SM, Holme I, Anderssen SA. Accelerometer-Determined Physical Activity in Adults and Older People. *Med Sci Sports Exerc*. 2012;44(2):266-72.

2. Healy GN, Wijndaele K, Dunstan DW, Shaw JE, Salmon J, Zimmet PZ, Owen N. Objectively measured sedentary time, physical activity, and metabolic risk. *Diabetes Care*. 2008;31(2):369-71.

3. WHO. Obesity: preventing and managing the global epidemic. Report of a WHO consultation. In: *Report of a WHO consultation*. WHO Technical Report Series 894, 2000:250.

4. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: Definitions and distinctions for health-related research. *Public Health Rep*. 1985;100(2):126-31.

5. Ziviani J, Scott J, Wadley D. Walking To School: Incidental Physical Activity in the Daily Occupations of Australian Children. *Occup Ther Int*. 2004;11(1):1-11.

6. Lubans DR, Hesketh K, Cliff DP, Barnett LM, Salmon J, Dollman J, Morgan PJ, Hills AP, Hardy LL. A systematic review of the validity and reliability of sedentary behaviour measures used with children and adolescents. *Obes Rev*. 2011;12(10):781-99.

7. Jerome GJ, Young DR, Laferriere DAN, Chen C, Vollmer WM. Reliability of RT3 Accelerometers among Overweight and Obese Adults. *Med Sci Sports Exerc*. 2009;41(1):110-4.

8. Dellava JE, Hoffman DJ. Validity of resting energy expenditure estimated by an activity monitor compared to indirect calorimetry. *Br J Nutr*. 2009;102(1):155-9.

9. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. *Med Sci Sports Exerc*. 2005;37(11 Suppl):S531-43.

10. Plasqui G, Westerterp KR. Physical activity assessment with accelerometers: an evaluation against doubly labeled water. *Obesity*. 2007;15(10):2371-9.

11. Ward DS, Evenson KR, Vaughn A, Rodgers AB, Troiano RP. Accelerometer Use in Physical Activity: Best Practices and Research Recommendations. *Med Sci Sports Exerc*. 2005;37(11):S582-S8.

12. Crouter SE, Clowers KG, Bassett DR, Jr. A novel method for using accelerometer data to predict energy expenditure. *J Appl Physiol*. 2006;100(4):1324-31.

13. Midorikawa T, Tanaka S, Kaneko K, Koizumi K, Ishikawa-Takata K, Futami J, Tabata I. Evaluation of low-intensity physical activity by triaxial accelerometry. *Obesity*. 2007;15(12):3031-8.

14. Bonomi AG, Goris AH, Yin BI, Westerterp KR. Detection of type, duration, and intensity of physical activity using an accelerometer. *Med Sci Sports Exerc*. 2009;41(9):1770-7.

15. Bassett DR, Jr. Validity and reliability issues in objective monitoring of physical activity. *Res Q Exerc Sport*. 2000;71(2 Suppl):S30-6.

16. Calabró MA, Welk GJ, Eisenmann JC. Validation of the SenseWear Pro Armband algorithms in children. *Med Sci Sports Exerc*. 2009;41(9):1714-20.

17. Abel MG, Hannon JC, Sell K, Lillie T, Conlin G, Anderson D. Validation of the Kenz Lifecorder EX and ActiGraph GT1M accelerometers for walking and running in adults. *Appl Physiol Nutr Metab*. 2008;33:1155-64.

18. Bouten CV, Koekkoek KT, Verduin M, Kodde R, Janssen JD. A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *IEEE Trans Biomed Eng*. 1997;44(3):136-47.

19. Slootmaker SM, Chin A Paw MJM, Schuit AJ, van Mechelen W, Koppes LLJ. Concurrent validity of the PAM accelerometer relative to the MTI Actigraph using oxygen consumption as a reference. *Scand J Med Sci Sports*. 2009;19:36-43.

20. Welk GJ, Blair SN, Wood K, Jones S, Thompson RW. A comparative evaluation of three accelerometry-based physical activity monitors. *Med Sci Sports Exerc*. 2000;32(9):S489-S97.

21. Rothney MP, Schaefer EV, Neumann MM, Choi L, Chen KY. Validity of physical activity intensity predictions by ActiGraph, Actical, and RT3 accelerometers. *Obesity*. 2008:16(8):1946-52.

22. Rothney MP, Apker GA, Song Y, Chen KY. Comparing the performance of three generations of ActiGraph accelerometers. *J Appl Physiol*. 2008;105(4):1091-7.

23. Baecke JA, Burema J, Frijters JE. A short questionaire for the measurement of habitual physical activity in epidemiological studies. *Am J Clin Nutr*. 1982;36:936-42.

24. Martin GS, Tapsell LC, Denmeade S, Batterham MJ. Relative validity of a diet history interview in an intervention trial manipulating dietary fat in the management of Type II diabetes mellitus. *Prev Med.*. 2003;36(4):420-8.

25. Schofield W. Predicting basal metabolic rate, new standards and review of previous work. *Hum Nutr Clin Nutr*. 1985;39(Supplement 1):5-41.

26. Fruin ML, Rankin JW. Validity of a Multi-Sensor Armband in Estimating Rest and Exercise Energy Expenditure. *Med Sci Sports Exerc*. 2004;36(6):1063-9.

27. Actigraph. *Actisoft Analysis Software 4.0 User's Manual*. Fort Walton Beach, FL: MTI Health Services, 2009;53.

28. Batterham M, Cavanagh R, Jenkins A, Tapsell L, Plasqui G, Clifton P. High-protein meals may benefit fat oxidation and energy expenditure in individuals with higher body fat. *Nutrition & Dietetics*. 2008;65(4):246-52.

29. Schoffelen PF, Westerterp KR, Saris WH, Ten Hoor F. A dual-respiration chamber system with automated calibration. *J Appl Physiol*. 1997;83(6):2064-72.

30. Weir JB. New methods for calculating metabolic rate with special reference to protein metabolism. *J Physiol*. 1949;109:1-9.

31. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. *Med Sci Sports Exerc*. 1998;30(5):777-81.

32. Kozey SL, Lyden K, Howe CA, Staudenmayer JW, Freedson PS. Accelerometer Output and MET Values of Common Physical Activities. *Med Sci Sports Exerc*. 2012;42(9):1776-84.

33. Staudenmayer J, Zhu W, Catellier DJ. Statistical Considerations in the Analysis of Accelerometry-Based Activity Monitor Data. *Med Sci Sports Exerc*. 2012;44(1S):S61-S7.

34. Crouter SE, Kuffel E, Haas JD, Frongillo EA, Bassett DR, Jr. A Refined 2-Regression Model for the ActiGraph Accelerometer. *Med Sci Sports Exerc*. 2010;42(5):1029-37.

35. Crouter SE et al., Validity of ActiGraph 2-Regression Model and Matthews and NHANES and Cut-Points for Assessing Free-Living Physical Activity. *J Phys Act Health. 2012;In Press.*

36. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307 - 10.

37. Schluter P. A multivariate hierarchical Bayesian approach to measuring agreement in repeated measurement method comparison studies. *BMC Med Res Methodol*. 2009;9(1):6.

38. Bland JM, Altman DG. Statistics Notes: Bayesians and Frequentists. Brit Med J. 1998;317:1151-1160.

39. Congdon P. *Bayesian Statistical Modelling* 2nd ed. West Sussex, UK: Wiley and Sons; 2007.

40. Congdon P. *Applied Bayesian Hierarchical Models*. Boca Raton, Fl: Chapman and Hall; 2010.

41. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res*. 1999;8:135 - 60.

42. Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: Evolution, critique and future directions. *Stat Med*. 2009;28(25):3049-67.

43. Fehling PC, Smith DL, Warner SE, Dalsky GP. Comparison of accelerometers with oxygen consumption in older adults during exercise. *Med Sci Sports Exerc*. 1999;31(1):171-5.

44. Campbell KL, Crocker PRE, McKenzie DC. Field evaluation of energy expenditure in women using Tritrac accelerometers. *Med Sci Sports Exerc*. 2002;34(10):1667-74.

45. Jakicic JM, Winters C, Lagally K, Ho J, Robertson RJ, Wing RR. The accuracy of the TriTrac-R3D accelerometer to estimate energy expenditure. *Med Sci Sports Exerc*. 1999;31(5):747-54.

46. Howe CA, Staudenmayer JW, Freedson PS. Accelerometer prediction of energy expenditure: Vector magnitude versus vertical axis. *Med Sci Sports Exerc*. 2009;41(12):2199-206.

47. Pettee Gabriel K, McClain JJ, High RR, Schmid KK, Whitfield GP, Ainsworth BE. Patterns of Accelerometer-Derived Estimates of Inactivity in Middle-Age Women. *Med Sci Sports Exerc*. 2012;44(1):104-10.

48. Kumahara H, Schutz Y, Ayabe M, Yoshioka M, Yoshitake Y, Shindo M, Ishii K, Tanaka H. The use of uniaxial accelerometry for the assessment of physical-activity-related energy expenditure: a validation study against whole-body indirect calorimetry. *Br J Nutr*. 2004;91(02):235-43.

49. Chen KY, Sun M. Improving energy expenditure estimation by using a triaxial accelerometer. *J Appl Physiol*. 1997;83(6):2112-22.

50. Chen KY, Bassett DRJ. The technology of accelerometry-based activity monitors: Current and future. *Med Sci Sports Exerc*. 2005;37(11):S490-S500.