Title: An evaluation of energy expenditure estimation by three activity monitors.

Running Title: Validity of three activity monitors.

Authors: Jennifer Ryan[1], John Gormley[1]

[1]Discipline of Physiotherapy, Trinity College Dublin

Corresponding author: Jennifer Ryan; Department of Physiotherapy, Trinity Centre for Health Sciences, St. James's Hospital, Dublin 8; phone: +35318963613; fax: +35314531915; email: ryanj8@tcd.ie

No funding was received for this work.

There is no conflict of interest for any author involved in this study.

Abstract

A comparative evaluation of the ability of activity monitors to predict energy expenditure (EE) is necessary to aid the investigation of the effect of EE on health.

Purpose: To validate and compare the RT3, the SWA and the IDEEA at measuring EE in adults and children. **Methods:** Twenty-six adults and 22 children completed a resting metabolic rate test and performed four treadmill activities (3 $km.h^{-1}$, 6 $km.h^{-1}$, 6 $km.h^{-1}$ at a 10% incline, 9 $km.h^{-1}$). Energy expenditure was assessed throughout the activities by the RT3, the SWA and the IDEEA. Indirect calorimetry was used as a criterion measure of EE against which each monitor was compared. Mean bias was assessed for by subtracting EE from IC from EE from each monitor for each activity. Limit of agreement plots were used to assess the agreement between each monitor and IC. **Results:** Limits of agreement for resting EE were narrowest for the RT3 for adults and children. Although the IDEEA displayed the smallest mean bias between measures at 3 $km.h^{-1}$, 6 $km.h^{-1}$, and 9 $km.h^{-1}$ in adults and children the SWA agreed closest with IC at 6 $km.h^{-1}$, 6 $km.h^{-1}$ at a 10% incline and 9 $km.h^{-1}$. Limits of agreement were closest for the SWA at 9 $km.h^{-1}$ in adults representing 42% of the overall mean EE. **Conclusions:** Although the RT3 provided the best estimate of resting EE in adults and children, the SWA provided the most accurate estimate of EE across a range of physical activity intensities.

Key Words: ACCELEROMETER, VALIDITY, ENERGY EXPENDITURE, PHYSICAL ACTIVITY, PHYSICAL INACTIVITY

Introduction

The positive effect of physical activity (PA) on health has been reported extensively (*US Department of Health and Human Services. 2008 Physical Activity Guidelines for Americans*, 2008). Yet, further investigation into the dose-response relationship between activity and health is needed. To accomplish this accurate and feasible measures of habitual PA are required. Although indirect measures such as questionnaires and interviews are often used to collect PA data in study populations (Ainsworth, 2009) they are limited by their inaccuracy (Bassett, Cureton, & Ainsworth, 2000; Rutten et al., 2003) and lack of precision. Criterion measures of energy expenditure (EE) such as indirect calorimetry (IC) and doubly-labeled water (DLW) are expensive and often unfeasible methods to use in large populations. Motion sensors may provide a feasible method of recording objective PA data.

Although considered motion sensors, pedometers are limited by their inablility to measure exercise volume or energy expenditure (Garber et al., 2011; Kumahara, Tanaka, & Schutz, 2009). Accelerometry-based devices, however, provide information on the frequency, duration and intensity of activity, allowing for comparison of data to PA guidelines. Traditional accelerometers measure the magnitude of the body's acceleration in one or more dimensions to provide an output in terms of 'counts'. Recent advances in monitoring technology have led to the development of advanced accelerometry-based devices that combine inputs from multiple sources and provide outputs in units other than counts. The Intelligent Device for Energy Expenditure and Activity (IDEEA) and the Sensewear Pro Armband (SWA) are two such devices. The SWA combines physiological data from multiple sensors to provide an output in terms of step count and EE. The IDEEA similarly combines accelerometry data from multiple sensors to provide information on gait parameters, posture and EE.

With the large number of accelerometry-based devices currently on the market, it is no longer sufficient to simply validate a device. A systematic approach should be implemented to prove that a new device provides an improvement in accuracy or feasibility over an already established measure. In order to compare accelerometry-based devices, a common outcome measure is required. Energy expenditure appears to be the most appropriate outcome measure to compare devices against as it is common to both traditional and advanced accelerometry-based devices. There are also criterion methods available to measure EE, such as IC and DLW.

Variations in experimental protocols, sample populations, data extraction methods and statistical analyses make it impossible to compare results across validation studies. A true comparative evaluation of monitors can only be provided when monitors are simultaneously compared to a criterion measure. The aim of this study was therefore to provide a comparative evaluation of the RT3 accelerometer, the SWA and the IDEEA by assessing their absolute and concurrent validity against IC.

Methods

Instrumentation

SWA

The SWA (Bodymedia, Inc.) is a lightweight (83 g) activity monitor that combines accelerometry data, heat loss data, skin temperature and galvanic skin response data to provide information on body position, step count and EE. The armband used in this study (Pro 2) obtains accelerometry data from a bi-axial accelerometer. Physiological data is incorporated with information regarding gender, age, height and weight to predict EE with the use of inbuilt algorithms. The manufacturers of the SWA periodically release software upgrades with new algorithms which they claim improve the estimation of EE. Participants wore the SWA on their right arm over the triceps muscle.

IDEEA

The IDEEA (Minisun LLC) consists of five sensors that are attached to the chest, anterior aspects of both thighs and soles of the feet, and a data collection device (59 g) that is clipped to the user's belt. The sensors contain biaxial accelerometers which collect data and transmit it through flexible wires to the recorder. Before each test the device is calibrated to ensure correct placement of the sensors. The IDEEA provides information regarding the type, duration, and, with the use of inbuilt proprietary algorithms, estimated energy cost of each activity carried out by the wearer.

RT3

The RT3 is a traditional accelerometer that measures accelerations in three dimensions (vertical, X; anteroposterior, Y; mediolateral, Z) with a triaxial accelerometer that is integrated into a single chip. It is a small, lightweight device (62.5 g) and provides an output in terms of 'counts'. Counts are converted into EE with the use of inbuilt proprietary algorithms to also provide a direct output of activity energy expenditure (AEE) and total energy expenditure (TEE). Participants wore the RT3 on their right hip throughout this study.

Indirect Calorimetry

Resting metabolic rate (RMR) and EE during the treadmill protocol were measured by indirect calorimetry (IC) using the Cosmed Quark (Rome, Italy). The Cosmed is a standard metabolic cart that measures ventilation and gas concentrations in expired air, using a flowmeter and gas analyser. The flowmeter consists of a bidirectional turbine and optoelectronic reader. The turbines were calibrated prior to each test session using a 3L syringe to ensure accurate volume measurements. Gas calibration, including room air calibration (assuming room air is 20.93% O_2 and 0.03% CO_2) and reference gas calibration (16% O_2, 5% CO_2), was conducted prior to each test.

A pilot study was conducted on 13 participants to assess the validity of the Cosmed. Participants lay supine while oxygen consumption (VO_2) was measured with the Cosmed and a douglas bag, for 10mins each, in a random order. The mean difference between the two methods was 0.21 $ml.kg^{-1}.min^{-1}$. Results of a paired t-test indicated there was no difference in VO_2 between methods (p=0.08). The 95% confidence intervals and Bland and Altman limits of agreement lay within ±2 $ml.kg^{-1}.min^{-1}$ (-0.03 $ml.kg^{-1}.min^{-1}$ to 0.45 $ml.kg^{-1}.min^{-1}$ and -0.57 $ml.kg^{-1}.min^{-1}$ to 0.99 $ml.kg^{-1}.min^{-1}$, respectively) indicating an acceptable level of agreement between the two methods (Atkinson, Davison, & Nevill, 2005).

Participants

The sample size required for this study was generated from data collected on 32 adults in our laboratory. Based on a mean difference of 0.15 $kcal.min^{-1}$ and a standard deviation of 0.216 $kcal.min^{-1}$, between the RT3 and IC at rest, 21 participants provided 80% power at the 0.05 α-level. Twenty-six adults (11 males, 15 females) and 22 children (11 males, 11 females) aged 6 to 36 years were recruited through the Faculty of Health Sciences and local schools. Ethical approval for this study was granted by the Faculty Ethics Committee.

The procedures and risks involved in the study were fully explained to participants and their guardians, where appropriate, before written informed consent was obtained.

Experimental Protocols

Height was measured to the nearest 0.5 cm with a calibrated stadiometer (SECA). Weight (to the nearest 0.1 kg), BMI and body fat percentage were measured in bare feet and light clothing using the Multi-Frequency Body Composition Analyser MC-180MA (Tanita Corp, Tokyo).

Participants attended the laboratory in the morning, at least 12 hrs post-prandial, and having refrained from caffeine, alcohol and vigorous exercise for 12 hrs prior to the test. They were also asked to refrain from nicotine and moderate exercise for 2 hrs prior to the test. A note was made of any recently taken medications (Compher, Frankenfield, Keim, & Roth-Yousey, 2006). Participants rested in a supine position for a minimum of 7 min while the activity monitors were initiated to record EE data every minute according to the manufacturer's specifications. The three activity monitors were attached to the participants as described in the instrumentation section). Resting metabolic rate (RMR) was measured for a minimum of

15 min using a ventilated hood in a thermoneutral environment (20-25 °C) and in the absence of external stimuli (Compher et al., 2006).

Following the measurement of RMR participants were given a 5 min familiarisation period with the treadmill (Viasys LE 300 CE). They were then fitted with a soft flexible facemask that held the flowmeter. Each time the facemask was applied it was checked to ensure there was an effective seal around the mouth. Each participant performed four activities of 5 mins duration in a randomised order: 1) walking at 3 $km.h^{-1}$; 2) walking at 6 $km.h^{-1}$; 3) walking at 6 $km.h^{-1}$ on a 10% incline; 4) running at 9 $km.h^{-1}$. Each activity was separated by a 5 min rest period in a seated position during which they were allowed to breathe without the facemask and drink water only. Participants were instructed not to hold onto the safety rail during treadmill locomotion.

Data Processing

Data from the three monitors and IC was downloaded following completion of the experimental protocol. Innerview Research Software version 6.1 was used to estimate EE from the SWA. Weir's equation was used to calculate EE from IC (Weir, 1949). Data from the monitors was time synchronised with that from IC.

The initial 5 min of RMR data were discarded. Resting metabolic rate was calculated from a continuous 5 min period of steady state data within the remaining time (Compher et al., 2006). Steady state was defined as a variation of <10% in VO_2 and VCO_2 and <5% in respiratory quotient (Compher et al., 2006). The final two minutes of EE data ($kcal.min^{-1}$) from supine lying and each treadmill activity was used to validate the monitors. Data was examined visually to check for malfunctioning units, time synchronisation and abnormal ouputs before statistical analysis.

Statistical Analysis

Data is reported as means ± SD. Normal distribution of data was assessed using the Shapiro-Wilk W test. Limit of agreement plots were calculated to assess agreement between EE from each monitor ($EE_{MONITOR}$) and EE from IC (EE_{IC}) (Bland & Altman, 1986). Bias was defined as the difference between $EE_{MONITOR}$ and EE_{IC}. Pearson's product moment correlations or Spearman's rank correlations were performed between $EE_{MONITOR}$ AND EE_{IC} for each activity. Statistical significance was considered at a two-sided $p<0.05$. All analyses were conducted using Analyse-It for Microsoft Excel, version 2.26.

Results

Descriptive statistics for adults and children are provided in tables 1. Adults were relatively young and lean with only four males and one female classified as overweight. Two children were considered overweight and two were considered obese according to the 2007 WHO BMI z-score for children (Butte, Garza, & de Onis, 2007).

Steady state RMR data was not obtained on two adults and four children. Furthermore RMR tests were not performed on three children as they refused to fast prior to the test or were unable to tolerate the ventilated hood. RT3 malfunction resulted in data from 5 adults being discarded. The IDEEA failed to record data on 3 participants. Five children were unable to complete 5 min running at 9 km.h^{-1}. Listwise deletion procedures were therefore employed to maximise sample sizes resulting in sample sizes ranging from n=13 to n=26.

Mean EE recorded by IC and the monitors are reported in tables 2 and 3. All monitors overestimated the energy cost of inactivity in adults. Although the RT3 overestimated resting energy expenditure (REE) in children by 0.14 kcal.min^{-1} (15%) it displayed the narrowest limits of agreement (LOA) of the three monitors. In children, the IDEEA displayed a large overestimation of REE (70%), wide LOA (-116% to 256% of mean EE), and a poor correlation with IC (-0.52).

The IDEEA showed the smallest mean bias at 3km.h^{-1} (+0.8%), 6km.h^{-1} (-0.3%) and 9km.h^{-1} (+3%) in adults. The large standard deviation of the bias at 6 km. h^{-1} and 9 km.h^{-1}, (±1.29 and ±2.24 kcal.min^{-1}, respectively), however, resulted in wide LOA. The SWA, in fact, showed closest agreement with IC at 6 km.h^{-1}, 6 km.h^{-1} on an incline and 9 km.h^{-1}. The RT3 overestimated the energy cost of all activities except walking on an incline. This overestimation of EE appeared to decrease with increasing speed, from +84% at 3 km.h^{-1} to +38% at 9 kmh^{-1}.

Similar results emerged for children. Although the IDEEA showed the smallest mean bias at 3 km.h^{-1} (+0.3%), 6 km.h^{-1} (-5%) and 9 km.h^{-1} (-5%), a large standard deviation of the mean bias resulted in wide LOA. Limits were narrowest for the SWA at 6 km.h^{-1} and 9 km.h^{-1}. The RT3 also overestimated the energy cost of walking at 3 km.h^{-1} (2%), 6 km.h^{-1} (21%) and running at 9 km.h^{-1} (23%) in children. In contrast to adults the magnitude of the overestimation of EE increased as speed increased.

All monitors were unable to detect the energy cost of traversing a slope. The RT3, however, displayed the smallest bias for this activity. This may be due to its large overestimation of the energy cost of walking on level ground. Apart from the IDEEA at rest and at 6km.h^{-1} on an incline, correlation coefficients between the three monitors and IC were moderate to good (r=0.63 to 0.91; p<0.01). The lower limit of the 95% confidence interval was as low as 0.27, for the RT3 at 3 km.h^{-1}, however, indicating a questionable association between the monitors and IC. See table 4 for more information.

Discussion

This study provides a comparative evaluation of three monitors that provide direct outputs of EE in small epochs. Results indicate that the SWA provides the best estimate of EE in both adults and children. The SWA demonstrated the closest agreement with criterion measured EE for four out of five activities in adults and three out of five activities in children.

The field of accelerometer research focuses, to a large extent, on the use of counts to record PA data. Variation in methods of data collection, processing, filtering and scaling, however, means that counts cannot be compared across different models of accelerometers (Chen & Bassett, 2005). Counts are also a meaningless value unless converted into a more interpretable unit. Recent research indicates that EE, as opposed to time spent in activity, predicts health related outcomes (Garber et al., 2011). Reporting exercise volume in terms of EE may therefore provide data regarding the dose-response relationship between activity and health as well as providing a common unit against which both traditional accelerometers and newer accelerometry-based devices can be compared.

As reported in previous studies (Arvidsson, Fitch, Hudes, & Fleming, 2011; Arvidsson, Slinde, Larsson, & Hulthen, 2007), the SWA underestimated the energy cost of treadmill activities in children. Although the magnitude of the underestimation was not the same across speeds, results from a previous study (Arvidsson et al., 2011) and the current study agreed that the magnitude of the bias was independent of speed. Critically, the non-systematic effect of speed on the bias between methods indicates that the bias cannot be easily corrected for by applying a correction factor. Activities were performed in a random order to reduce any sources of unknown bias on results. It is not known, however, if this randomisation impacted the association between speed and bias.

Comparisons of the IDEEA and the SWA in children have revealed the IDEEA to provide a better estimation of EE, in terms of mean bias, during track and treadmill locomotion (Arvidsson et al., 2011; Arvidsson, Slinde, Larsson, & Hulthen, 2009). These studies did not report LOA, however. If conclusions from the current study were based on the mean bias alone, the IDEEA would be considered most accurate at 3 $km.h^{-1}$, 6 $km.h^{-1}$ and 9 $km.h^{-1}$. The LOA clearly show, however, that there was closer agreement between the SWA and IC at 6 $km.h^{-1}$ and 9 $km.h^{-1}$, demonstrating that bias alone does not accurately portray the level of agreement between measures. Sun et al. (Sun, Schmidt, & Teo-Koh, 2008) reported that the RT3 overestimated the energy cost of children walking at 3 $km.h^{-1}$ and 6 $km.h^{-1}$ by 54% and 96%, respectively. Although the RT3 also overestimated EE at these speeds in the current study, a smaller mean bias was found: +2% and +21%, respectively. Kavouras et al. (Kavouras, Sarras, Tsekouras, & Sidossis, 2008) and Hussey et al. (Hussey et al., 2009), however, reported that the RT3 underestimated EE at 6 $km.h^{-1}$. As LOA were not reported by any of these studies it is difficult to comment on their findings in relation to the data recorded during this study.

Variations in statistical analysis aside, contradictory results between studies may be due to differences in the age range of participants. Although the SWA appeared the most accurate measure in both adults and children, LOA found in this study were wider for children than adults for all activities. The accuracy of the proprietary equations depends on the sample population in which they are developed. Not only may equations developed on adults not be applicable to children, changes in children's walking patterns with age (Bjornson et al., 2011) may reduce the accuracy of the monitors across age. The RT3, IDEEA and SWA have been validated in children aged 7-14 yr, 8-17 yr, and 8-15 yr, respectively (Arvidsson et al., 2011; Arvidsson, Slinde, & Hulthen, 2009; Arvidsson et al., 2007; Arvidsson, Slinde, Larsson, et al., 2009; Calabro, Welk, & Eisenmann, 2009; Dorminy, Choi, Akohoue, Chen, & Buchowski, 2008; Hussey et al., 2009; Kavouras et al., 2008; Sun et al., 2008; Zhang, Pi-Sunyer, & Boozer, 2004). This study provides a unique evaluation of these monitors in children aged 6-17 years.

To date validation of the SWA (v6.1) during locomotion has only been conducted on endurance athletes (Drenowatz & Eisenmann, 2011; Koehler et al., 2011). Although the participants in the current study may be more representative of the general population, the results of the current study indicate that, as reported for endurance athletes, the SWA underestimates the energy cost of running. It has been suggested that the SWA's estimation of EE plateaus at an intensity equivalent of 10METs (Drenowatz & Eisenmann, 2011). When adults in the current study were running at $9km.h^{-1}$ they were exercising at a mean intensity of 11.68METs. Despite this, of all the activities, the SWA agreed best with IC at $9km.h^{-1}$ (LOA: -29% to 13%), suggesting that exercising above 10METs did not decrease the accuracy of the SWA in this study. Although a newer version of the SWA software (v7.0) has been released since this study was conducted the manufacturers indicate that the algorithms included in v7.0 do not differ from those in v6.1 (written communication).

The three monitors assessed in this study provide an estimation of total energy expenditure (TEE) which includes a REE component. As REE can account for up to 60% of TEE (Levine, 2005), if a monitor is unable to accurately predict REE it is likely that its measure of TEE will also be inaccurate. This is the first study to report on the ability of the RT3 to measure REE in adults and children. Only one study has evaluated the IDEEA as a measure of REE and although only a small overestimation of EE by the IDEEA was reported (10%) LOA were not calculated (Arvidsson, Slinde, Larsson, et al., 2009). The RT3 was the most accurate at measuring REE in adults and children. Despite this EE was overestimated by 29% and 15% with LOA of 8% to 49% and -8% to 37% in adults and children, respectively. The large overestimation of REE, wide LOA (-116% to 256% of mean EE), and poor correlation (-0.52) observed for IDEEA data on children may have been caused by three extreme recordings for REE (0 $kcal.min^{-1}$, 2.33 $kcal.min^{-1}$ and 3.59 $kcal.min^{-1}$). As there was no explanation for these values and data recorded from the same children for the remaining activities was not extreme, the removal of these children from data analysis was not justified. The decision not to remove this data from analyses was supported by Arvidsson et al. (Arvidsson, Slinde, & Hulthen, 2009) who also reported that the IDEEA gave extreme values for REE in children. Despite the small bias between the IDEEA and IC for many activities, the IDEEA needs to be consistently accurate for it to be considered an acceptable method of measuring EE. Monitor malfunction and non-compliance, although anecdotally common, are often underreported in studies. It is vital that information on the feasibility of monitors is included in reports as it is an important consideration in monitor selection (Trost, McIver, & Pate, 2005).

Of the three monitors the IDEEA took the longest to initiate because of the calibration required. Some participants, particularly children, experienced discomfort from the foot sensors of the IDEEA. The IDEEA is also the most difficult of the monitors to attach. The difficulty involved in reattaching the monitor following washing may reduce compliance with it. There was no data missing for the analysis of the SWA. The armband occasionally became loose, however, and had to be refitted. Participants generally remarked that the SWA and the RT3 were comfortable and indiscrete.

This study is limited by the lack of free-living activities in the protocol. The time commitment of participants involved in this study limited our ability to include these in the study. Locomotor activity, however, is the predominant activity in a person's day. Therefore the validation of accelerometers during this activity is of primary importance (Welk, 2005). The accuracy of a monitor should also be assessed across a range of PA intensities, including inactivity (Matthew, 2005), all of which were captured in this study. Future studies should comparatively evaluate these monitors during free-living activities.

A further limitation of this study was the comparison of EE during steady-state exercise. Adults and children require 3-6mins to reach a steady-state work rate (Turley & Wilmore, 1997). Habitual PA is usually accrued in a sporadic manner, however, particularly by children (Mark & Janssen, 2009). Reliance on steady-state data means that the ability of activity monitors to accurately record the energy cost of non-steady-state exercise during short bursts of activity was not tested. Also, as heart-rate was not monitored during the 5min rest periods it is not known if participants fully recovered between activities. Reports suggest, however, that adults and children recover from submaximal exercise in 3-5 min (Turley & Wilmore, 1997). Finally, the authors acknowledge that the sample sizes in this study are small. A power calculation indicated that a sample size of 21 provided 80% power at a 0.05 α-level, however. The sample size in the current study is also in line with the sample sizes reported in previous validation studies of these three monitors (Arvidsson et al., 2007; Calabro et al., 2009; Drenowatz & Eisenmann, 2011; Johannsen et al., 2010)

In conclusion, many factors are involved in selecting an activity monitor including the research hypothesis, feasibility, validity and the burden on participants and researchers (Welk, 2005). This comparative evaluation reveals that the SWA appears to be the most accurate monitor to use for the assessment of EE in adults and children. It also appears to be a feasible method of measuring habitual PA with little or no monitor malfunction or participant discomfort reported.

Acknowledgements

The authors thank the participants and their parents who volunteered their time for this study.
This study was supported by a research grant from Trinity College Dublin.
There is no conflict of interest.

References:

Ainsworth, B. E. (2009). How do I measure physical activity in my patients? Questionnaires and objective methods. *British Journal of Sports Medicine, 43*(1), 6-9. doi: bjsm.2008.052449 [pii] 10.1136/bjsm.2008.052449 [doi]

Arvidsson, D., Fitch, M., Hudes, M. L., & Fleming, S. E. (2011). Accuracy of multisensor activity monitors in normal versus high BMI African American children. *Journal of Physical Activity and Health, 8*(8), 1124-1134.

Arvidsson, D., Slinde, F., & Hulthen, L. (2009). Free-living energy expenditure in children using multi-sensor activity monitors. *Clinical Nutrition, 28*(3), 305-312.

Arvidsson, D., Slinde, F., Larsson, S., & Hulthen, L. (2007). Energy cost of physical activities in children: validation of SenseWear Armband. *Medicine and science in sports and exercise, 39*(11), 2076-2084. doi: 10.1249/mss.0b013e31814fb439 [doi]

00005768-200711000-00025 [pii]

Arvidsson, D., Slinde, F., Larsson, S., & Hulthen, L. (2009). Energy cost in children assessed by multisensor activity monitors. *Medicine and science in sports and exercise, 41*(3), 603-611. doi: 10.1249/MSS.0b013e31818896f4 [doi]

Atkinson, G., Davison, R. C., & Nevill, A. M. (2005). Performance characteristics of gas analysis systems: what we know and what we need to know. *International Journal of Sports Medicine, 26 Suppl 1*, S2-10. doi: 10.1055/s-2004-830505 [doi]

Bassett, D. R., Jr., Cureton, A. L., & Ainsworth, B. E. (2000). Measurement of daily walking distance-questionnaire versus pedometer. *Medicine and science in sports and exercise, 32*(5), 1018-1023.

Bjornson, K. F., Song, K., Zhou, C., Coleman, K., Myaing, M., & Robinson, S. L. (2011). Walking stride rate patterns in children and youth. *Pediatric Physical Therapy, 23*(4), 354-363. doi: 10.1097/PEP.0b013e3182352201

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet, 1*(8476), 307-310.

Butte, N. F., Garza, C., & de Onis, M. (2007). Evaluation of the feasibility of international growth standards for school-aged children and adolescents. *Journal of Nutrition, 137*(1), 153-157.

Calabro, M. A., Welk, G. J., & Eisenmann, J. C. (2009). Validation of the SenseWear Pro Armband algorithms in children. *Medicine and science in sports and exercise, 41*(9), 1714-1720. doi: 10.1249/MSS.0b013e3181a071cf [doi]

Chen, K. Y., & Bassett, D. R., Jr. (2005). The technology of accelerometry-based activity monitors: current and future. *Medicine and science in sports and exercise, 37*(11 Suppl), S490-500. doi: 00005768-200511001-00002 [pii]

Compher, C., Frankenfield, D., Keim, N., & Roth-Yousey, L. (2006). Best practice methods to apply to measurement of resting metabolic rate in adults: a systematic review. *Journal of American Dietetic Association, 106*(6), 881-903. doi: 10.1016/j.jada.2006.02.009

Dorminy, C. A., Choi, L., Akohoue, S. A., Chen, K. Y., & Buchowski, M. S. (2008). Validity of a multisensor armband in estimating 24-h energy expenditure in children. *Medicine and science in sports and exercise, 40*(4), 699-706. doi: 10.1249/MSS.0b013e318161ea8f [doi]

Drenowatz, C., & Eisenmann, J. C. (2011). Validation of the SenseWear Armband at high intensity exercise. *European Journal of Applied Physiology, 111*(5), 883-887. doi: 10.1007/s00421-010-1695-0 [doi]

Garber, C. E., Blissmer, B., Deschenes, M. R., Franklin, B. A., Lamonte, M. J., Lee, I. M., . . . American College of Sports, M. (2011). American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Medicine and science in sports and exercise, 43*(7), 1334-1359.

Hussey, J., Bennett, K., Dwyer, J. O., Langford, S., Bell, C., & Gormley, J. (2009). Validation of the RT3 in the measurement of physical activity in children. *Journal of Science and Medicine in Sport, 12*(1), 130-133. doi: S1440-2440(07)00219-8 [pii]

10.1016/j.jsams.2007.09.010 [doi]

Johannsen, D. L., Calabro, M. A., Stewart, J., Franke, W., Rood, J. C., & Welk, G. J. (2010). Accuracy of armband monitors for measuring daily energy expenditure in healthy adults. *Medicine and science in sports and exercise, 42*(11), 2134-2140. doi: 10.1249/MSS.0b013e3181e0b3ff [doi]

Kavouras, S., Sarras, S., Tsekouras, Y., & Sidossis, L. (2008). Assessment of energy expenditure in children using the RT3 accelerometer. *Journal of Sports Sciences, 26*(9), 959-966.

Koehler, K., Braun, H., M, D. E. M., Fusch, G., Fusch, C., & Schaenzer, W. (2011). Assessing Energy Expenditure in Male Endurance Athletes: Validity of the SenseWear Armband. *Medicine and Science in Sports and Exercise, 43*(7), 1328-1333. doi: 10.1249/MSS.0b013e31820750f5 [doi]

Kumahara, H., Tanaka, H., & Schutz, Y. (2009). Are pedometers adequate instruments for assessing energy expenditure? *Europena Journal of Clinical Nutrition, 63*(12), 1425-1432. doi: ejcn2009108 [pii]

10.1038/ejcn.2009.108 [doi]

Levine, J. A. (2005). Measurement of energy expenditure. *Public Health and Nutrition, 8*(7A), 1123-1132. doi: S1368980005001382 [pii]

Mark, A. E., & Janssen, I. (2009). Influence of bouts of physical activity on overweight in youth. *American Journal of Preventative Medicine, 36*(5), 416-421. doi: S0749-3797(09)00094-4 [pii]

10.1016/j.amepre.2009.01.027 [doi]

Matthew, C. E. (2005). Calibration of accelerometer output for adults. *Medicine and science in sports and exercise, 37*(11 Suppl), S512-522.

Rutten, A., Ziemainz, H., Schena, F., Stahl, T., Stiggelbout, M., Auweele, Y. V., . . . Welshman, J. (2003). Using different physical activity measurements in eight European countries. Results of the European Physical Activity Surveillance System (EUPASS) time series survey. *Public Health and Nutrition, 6*(4), 371-376. doi: 10.1079/PHN2002450 [doi]

Sun, D. X., Schmidt, G., & Teo-Koh, S. M. (2008). Validation of the RT3 accelerometer for measuring physical activity of children in simulated free-living conditions. *Pediatric Exercise Science*, *20*(2), 181-197.

Trost, S. G., McIver, K. L., & Pate, R. R. (2005). Conducting accelerometer-based activity assessments in field-based research. *Medicine and science in sports and exercise, 37*(11 Suppl), S531-543.

Turley, K. R., & Wilmore, J. H. (1997). Cardiovascular responses to treadmill and cycle ergometer exercise in children and adults. *Journal of Applied Physiology*, *83*(3), 948-957.

. *US Department of Health and Human Services. 2008 Physical Activity Guidelines for Americans*. (2008). Washington, DC: www.health.gov/paguidelines.

Weir, J. B. (1949). New methods for calculating metabolic rate with special reference to protein metabolism. *Journal of Physiology, 109*(1-2), 1-9.

Welk, G. J. (2005). Principles of design and analyses for the calibration of accelerometry-based activity monitors. *Medicine and science in sports and exercise, 37*(11 Suppl), S501-511.

Zhang, K., Pi-Sunyer, F. X., & Boozer, C. N. (2004). Improving energy expenditure estimation for physical activity. *Medicine and science in sports and exercise*

36(5), 883-889. doi: 00005768-200405000-00022 [pii]

Table 1. Descriptive statistics for physical characteristics of adults and children

	Adults (n=26; 11men)	Children (n=22; 11 boys)
Age (yr)	24.7±4.4	11.5±3.0
Weight (kg)	69.5±12.0	44.9±13.9
Height (cm)	174.3±8.5	153.9±16.4
BMI ($kg.m^{-2}$)	22.8±2.9	18.4±3.0
Body Fat (%)	22.0±6.3	22.0±6.3
RMR ($kcal.d^{-1}$)	1397.2±283.6	1308.6±194.6

Table 2. Mean energy cost of rest and treadmill activities in adults as measured by IC and the three monitors.

	IC (kcal.min^{-1}) (n=24 to 26)	SWA (kcal.min^{-1}) (n=26)	RT3 (kcal.min^{-1}) (n=21)	IDEEA (kcal.min^{-1}) (n=25)
Rest	0.97±0.22	1.11±0.18	1.26±0.19	1.28±0.18
3 km.h^{-1}	3.91±0.62	4.48±0.94	4.36±1.03	3.94±0.67
6 km/h^{-1}	6.28±0.93	6.19±1.18	8.99±2.20	6.26±1.51
6 km.h^{-1}@10% incline	11.47±2.12	6.86±1.41	8.87±2.27	6.07±1.50
9 km.h^{-1}	11.41±2.31	10.54±2.07	15.94±3.41	11.71±2.63

Table 3. Mean energy cost of rest and treadmill activities in children as measured by IC and the three monitors.

	IC (kcal.min^{-1}) (n=15 to 22)	SWA (kcal.min^{-1}) (n=17 to 22)	RT3 (kcal.min^{-1}) (n=17 to 22)	IDEEA (kcal.min^{-1}) (n=15 to 20)
Rest	0.91±0.14	0.79±0.23	1.01±0.17	1.60±0.71
3 km.h^{-1}	3.22±0.67	3.21±1.59	3.30±0.93	3.22±0.45
6 km/h^{-1}	5.37±1.28	4.73±1.69	6.49±2.11	5.14±1.41
6 km.h^{-1}@10% incline	7.91±2.56	5.00±1.72	6.44±1.95	4.73±1.06
9 km.h^{-1}	9.36±2.48	8.46±2.83	11.55±4.10	8.98±1.15

Table 4 Limits of agreement and correlations between IC and the SWA, the RT3 and the IDEEA for the assessment of rest and treadmill activities [Bias and LOA in $kcal.min^{-1}$].

		Adults			Children		
		Bias	LOA	Correlation Coefficient [95% CI]	Bias	LOA	Correlation coefficient [95% CI]
Rest	SWA	0.14	-0.07 to 0.35	0.86* [0.71 to 0.94]	-0.07	-0.34 to 0.21	0.71* [0.32 to 0.90]
	RT3	0.28	0.08 to 0.48	0.87* [0.68 to 0.95]	0.14	-0.07 to 0.34	0.78* [0.45 to 0.92]
	IDEEA	0.31	0.11 to 0.51	0.91* [0.79 to 0.96]	0.64	-1.06 to 2.33	-0.52 [-0.83 to 0.04]
$3km.h^{-1}$	SWA	0.56	-0.68 to 1.81	0.72* [0.46 to 0.86]	-0.01	-2.32 to 2.29	0.75* [0.47 to 0.89]
	RT3	3.30	1.39 to 5.22	0.63* [0.27 to 0.84]	0.07	-0.91 to 1.06	0.85* [0.68 to 0.94]
	IDEEA	0.03	-0.82 to 0.88	0.74* [0.48 to 0.88]	0.01	-0.88 to 0.89	0.78* [0.52 to 0.91]
$6km.h^{-1}$	SWA	-0.10	-1.62 to	0.78*	-0.64	-2.41 to	0.82*

			1.43	[0.57 to 0.91]		1.13	[0.62 to 0.92]
	RT3	2.66	-1.06 to 6.39	0.69* [0.37 to 0.87]	1.12	-1.56 to 3.81	0.67* {0.34 to 0.84]
	IDEEA	-0.02	-2.56 to 2.51	0.64* [0.33 to 0.83]	-0.28	-2.15 to 1.59	0.75* [0.46 to 0.90]
$6km.h^{-1}$ @ 10%	SWA	-4.60	-7.36 to -1.84	0.79* [0.58 to 0.90]	-2.91	-6.18 to 0.37	0.71* [0.42 to 0.87]
	RT3	-2.74	-6.57 to 1.09	0.66* [0.33 to 0.85]	-1.48	-5.13 to 2.18	0.67* [0.35 to 0.85]
	IDEEA	-5.29	-8.65 to -1.94	0.74* [0.49 to 0.88]	-3.16	-7.29 to 0.98	0.56 [0.16 to 0.80]
$9km.h^{-1}$	SWA	-0.87	-3.27 to 1.54	0.87* [0.73 to 0.94]	-0.91	-3.62 to 1.80	0.80* [0.53 to 0.93]
	RT3	4.35	-1.02 to 9.72	0.68* [0.36 to 0.86]	2.18	-2.33 to 6.70	0.83* [0.59 to 0.94]
	IDEEA	0.33	-4.06 to 4.71	0.67* [0.37 to	-0.48	-4.64 to 3.69	0.69* [0.27 to

	0.84]	0.89]

*$p<0.01$; Sample sizes range from n=13 to n=26 for all activities

Abbreviations: LOA = limits of agreement; 95% CI = 95% confidence intervals