

2011

Physical activity assessment under free-living conditions using pattern-recognition monitors

Miguel Andres Calabro
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/etd>

 Part of the [Kinesiology Commons](#)

Recommended Citation

Calabro, Miguel Andres, "Physical activity assessment under free-living conditions using pattern-recognition monitors" (2011).
Graduate Theses and Dissertations. 11998.
<https://lib.dr.iastate.edu/etd/11998>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Physical activity assessment under free-living conditions using pattern recognition monitors

by

Miguel Andrés Calabró

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Kinesiology (Behavioral Basis of Physical Activity)

Program of Study Committee:
Gregory J. Welk, Major Professor
Warren D. Franke
Amy S. Welch
Alicia Carriquiry
Lorraine Lanningham-Foster

Iowa State University

Ames, Iowa

2011

Copyright © Miguel Andrés Calabró, 2011. All rights reserved.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	vi
ABSTRACT.....	vii
CHAPTER 1. INTRODUCTION	1
Review of Past Research.....	3
Comparison of the SenseWear Armband with a second pattern-recognition monitor and a self-report instrument	3
Evaluation of the SWA in free-living adults.....	3
Evaluation of the SWA monitor in older adults.....	4
Evaluation of the SWA in children.....	4
Evaluation of the SWA in Adults with DLW	5
Evaluation of the SWA during lower intensity activities	6
Overview and Purpose of Dissertation Research.....	6
CHAPTER 2. EXTENDED REVIEW OF LITERATURE: OBJECTIVE ACTIVITY MONITORING UNDER FREE-LIVING CONDITIONS.....	8
Accelerometry-based activity monitors	9
Pattern-recognition monitors	12
References.....	25
CHAPTER 3. PHYSICAL ACTIVITY ASSESSMENT IN OLDER ADULTS UNDER FREE-LIVING CONDITIONS.	38
Abstract.....	38
Introduction.....	40

Methods.....	42
Participants.....	42
Physical Activity Assessment methods	43
Data collection procedures.....	45
Data Processing.....	47
Statistical analyses	49
Results.....	51
Descriptive Analyses	51
Analyses of Total energy expenditure	52
Analyses of Activity energy expenditure.....	53
Discussion.....	54
References.....	63
Tables.....	70
Table 1. Representation of the experimental protocol.....	70
Table 2. Descriptive statistics for the sample population.....	71
Table 3. Total daily energy expenditure values (kcal.day ⁻¹) for each assessment tool...	72
Figures.....	73
Figure 1- Moderate and vigorous physical activity contribution by assessment tool (minutes).....	73
Figure 2- Bland Altman plots for total energy expenditure (kcal.day ⁻¹).....	74
Figure 3- Total energy expenditure distribution by assessment tool.....	76
Figure 4- Bland Altman plots for activity energy expenditure (AEE, kcal.day ⁻¹).....	77

CHAPTER 4. VALIDATION OF PATTERN-RECOGNITION MONITORS IN CHILDREN USING DOUBLY LABELED WATER.....	79
Abstract.....	79
Introduction.....	81
Methods.....	83
Instruments.....	83
Data Collection Procedures.....	84
Statistical analyses	88
Results.....	89
Total energy expenditure comparisons	90
Activity energy expenditure.....	92
Discussion.....	93
Tables	99
Table 1. Representation of the experimental protocol.....	99
Table 2. Descriptive statistics.	100
Table 3. Total energy expenditure (TEE) values from the activity monitors and DLW method	101
Figures.....	102
Figure 1- Bland-Altman plots for total energy expenditure (kcal·day ⁻¹).	102
Figure 2- Bland-Altman plots for activity energy expenditure (kcal·day ⁻¹).	104
CHAPTER 5. SUMMARY.....	110
References.....	113

APPENDIX A.....	115
Validation of the SenseWear Pro Armband algorithms in children	116
Accuracy of armband monitors for measuring daily energy expenditure in healthy adults.....	138
APPENDIX B	164
Figure 1- New SenseWear Armband algorithms (5.2)	164

ACKNOWLEDGEMENTS

Many people have contributed to the completion of my dissertation:

- First I would like to thank Dr. Greg Welk for his mentorship and support during my years as a graduate student, and for his many contributions to this dissertation.
- I would like to acknowledge my committee members and professors for their guidance throughout my Ph.D. program.
- I would like to acknowledge the enthusiastic support of many volunteers that participated in my studies over the years.
- I would like to thank my fellow graduate students and the whole Kinesiology Department at ISU, for their help and support during my years as a graduate student.
- I would like to thank my friends for their support over the years.
- I am very grateful to my family, to whom I will be forever indebted for all of their love and support.
- Lastly, I would like to express my infinite gratitude to my wife Gaby for her continues love and support over the years.

ABSTRACT

Extensive literature has documented the health benefits of physical activity. Valid, reliable and feasible physical activity assessment tools are necessary to assess the complexity and multidimensionality of physical activity behavior. Pattern-recognition activity monitors that integrate information from multiple sensors appear to be the most promising approach for assessing physical activity under free-living conditions. Previous studies have provided support to the validity of pattern-recognition monitors for assessing the energy cost of activity under-free living conditions in young adults. However, children and older adults present unique measurement challenges for the assessment of physical activity under free-living conditions. The series of studies in this dissertation extends previous research by assessing the accuracy of a pattern-recognition monitor (SenseWear Armband) in children and older adults under free-living conditions. Consistent with previous findings in young adults, results indicate that the SenseWear Armband monitors provide valid estimates of total energy expenditure and activity energy expenditure in older adults and children, under free-living conditions. Collectively, the findings of this research support the validity of the SenseWear Armband for assessing physical activity under free-living conditions in children and older adults.

CHAPTER 1. INTRODUCTION

The importance of valid, reliable and feasible physical activity assessment tools for the assessment of physical activity (PA) under free-living conditions has been well established (Troiano 2005, Welk 2002). The development of more accurate measurement tools would help to advance research in a number of ways: evaluating associations between PA and health outcomes, understanding factors that influence PA behavior, capturing changes in behavior following PA interventions, and exploring trends and patterns of PA in the population. The accurate characterization of the pattern, duration, and intensity of PA is essential for understanding activity behavior and for effective disease prevention programming.

The ability of pattern-recognition monitors to assess physical activity under-free living conditions has been one of our main areas of interest. Pattern-recognition monitors integrate data from multiple sensors and use algorithms to estimate energy expenditure (EE), allowing for more specific PA assessment. The SenseWear Pro Armband (SWA, Bodymedia Inc., Pittsburgh, PA) is one of the most promising pattern recognition monitors as it has good validity (i.e., it provides accurate information about physical activity and energy expenditure) as well as good utility (i.e., it is easy and comfortable to wear).

The SWA combines data from accelerometers with information from several heat related sensors (Jakicic et al. 2004). A unique feature of the SWA is its ability to measure heat dissipation, in combination with acceleration. This provides considerable advantages over other accelerometry-based activity monitors that only assess movement. In

particular, the ability to detect heat production provides additional information about the intensity and metabolic cost of PA. A number of studies have supported the validity of the SWA for use in adults (Fruin et al. 2004; Jakicic et al. 2004; King et al. 2004; Malavolti et al. 2006; Cereda et al. 2007) and children (Arvidsson et al. 2007; Dorminy et al. 2008; Calabro et al. 2009b) but most studies have been conducted under controlled laboratory conditions. Studies under free-living conditions provide a more robust test of validity since the nature and variety of activities performed are more complex.

A unique aspect of the SWA is that the proprietary algorithms applied by the software to estimate EE are continuously being improved by the manufacturer in an effort to reduce error. Several studies have demonstrated that the enhancements improve the accuracy of the assessment (Jakicic, et al., 2004; Calabro et al., 2009). Our laboratory has contributed to this work by collecting data needed to “train” the algorithms. Our lab has also conducted several recent studies that have evaluated the validity of the revised SWA under field conditions.

This dissertation builds upon this line of research and, provides the best available test of the validity of the SWA monitors (and associated algorithms) for assessing EE under free-living conditions. The subsequent sections describe our past research and provide a justification for my dissertation research.

Review of Past Research

Comparison of the SenseWear Armband with a second pattern-recognition monitor and a self-report instrument

In an early study conducted in 2006, we compared the SWA (Software V. 4.1) with a second pattern-recognition monitor (Intelligent Device for Energy Expenditure and Activity, IDEEA) and a PA recall instrument (24PAR) in middle age adults (mean age: 29.9 ± 5.7 years). The results showed that the SWA yielded similar estimates of total energy expenditure (TEE) and minutes of PA as the more established IDEEA monitor, despite being less invasive and considerably less expensive. In addition, the results supported the validity of the recall-based, self-report measure, compared to both pattern recognition monitors. This study was published in *Journal of Physical Activity and Health* (Calabro et al. 2009a).

Evaluation of the SWA in free-living adults

In another study in our laboratory conducted in 2007, we tested the accuracy of the SWA (Software V. 5.1) assessing complex recreational and lifestyle activities (i.e., walking, computer work, tennis, stationary biking, step stepping and leisure time) in a sample of adults (mean age= 27.6 ± 4.4 years). Participants wore the SWA and had oxygen consumption measured using a portable metabolic analyzer. Minute by minute EE correlations varied across participants but the mean correlation ($r = 0.69$) demonstrated good correspondence between the measures. There were significant differences in EE

estimates for biking ($t = 5.85$, $p < .01$) but not for tennis, box stepping, computer work or the leisure time. These data were shared with the company to assist in algorithm development.

Evaluation of the SWA monitor in older adults

In a subsequent study conducted in 2007, we assessed the validity of the SWA (Software V. 5.1) in a sample of 26 older adults (mean age = 74.1 ± 4.0) conducting tasks associated with usual activities of daily living. On average, the SWA overestimated EE during sitting (7.0%), fidgeting while sitting (1.7%), dressing (49.8%), stocking groceries (45.5%), stair stepping (6.1%), laundry folding (26.5%), sweeping (21.2%) and walking (23.7%), and underestimated EE during sitting and standing (5.2%) compared to indirect calorimetry. Even though some of the estimations were significantly different, the results demonstrated good temporal agreement as the patterns from the monitor mirrored the actual changes in EE across the stages. Furthermore, the SWA demonstrated a unique ability to detect subtle changes in EE associated with lower intensity household tasks in older adults. These data were shared with the company to assist in algorithm development.

Evaluation of the SWA in children

In 2008, we conducted a formal evaluation of the new proprietary algorithms (software V. 6.1) that were developed specifically from children's data. Twenty one healthy children, averaging $9.4 (\pm 1.3)$ years of age, participated in different activities

while being monitored with the SWA and a stationary metabolic analyzer. The activity protocol lasted 41 minutes and included: resting (lying down), coloring (sitting), playing computer games, walking on a treadmill (2, 2.5 and 3mph) and stationary biking. With the original algorithms (software V.5.1), the SWA was found to overestimate EE by 32%. Results with the newly developed algorithms yielded non-significant differences in overall estimates of EE across the 41 minute trial (error = 1.7%). In addition, only two activities (resting in a supine position and biking) yielded EE estimations that were significantly different from the criterion measure ($p < 0.001$). The average absolute difference in EE estimates for the various activities was 13% and the average individual correlation (conducted across the 41 minute-trial) was 0.70. These results provided good support for the validity of the newly developed SWA algorithms to estimate PAEE in children. This study was published in *Medicine and Science in Sports and Exercise* (Calabro et al. 2009b), see Appendix A.

Evaluation of the SWA in Adults with DLW

In 2009, we conducted a formal evaluation of the accuracy of the SWA (including the latest version of the software (V.6.1)), and the newest version of the monitor (SWA mini (SenseWear Software 7.0)) against criterion data from the doubly-labeled water (DLW) technique in a sample of 30 adults. Absolute error rates of ~8.0% were observed for the two versions of the SWA. The results of the study provided the best available evidence to support the accuracy of the SWA and the Mini monitors for measuring EE

under free-living conditions. This study was published in *Medicine and Science in Sports and Exercise* (Johannsen et al. 2010), see appendix A.

Evaluation of the SWA during lower intensity activities

In 2010, we evaluated the relative ability of different PA assessment tools for capturing lower intensity activities in young adults. In this study we compared data from 3 pattern-recognition monitors (SWA, v. 6.1, Mini, v. 7.0 and Actiheart monitor), and 2 accelerometry-based activity monitors (ActiGraph GT3X and ActivPAL) to a portable metabolic analyzer (Oxycon Mobile) during 60 minutes of free-living activity. The results showed that the Mini and SWA monitors provided more accurate estimates of EE during light to moderate intensity free-living activities (9.5-14.1% absolute error), compared to other activity monitors (23.6-30.5% absolute error). The results of this study showed the relative validity of the SWA compared to other selected PA assessment tools. This study remains in preparation (*Personal communication*, April 2011).

Overview and Purpose of Dissertation Research

This previous research helped to advance the knowledge base regarding the accuracy of the SWA under different conditions and with different samples. However, the studies also identified additional types of validation research that were needed to further evaluate these monitors. Two populations that merited additional research were young children and older adults. Young children present unique measurement challenges due to the inherent complexities associated with variability in growth and maturation. Older

adults present unique measurement challenges due to inherent variability in health, cognition and functional capacity. The previous work in our laboratory, and by others, has been conducted with both populations under laboratory conditions so additional work was needed to evaluate validity under free-living conditions. The research presented in this dissertation builds from our previous research by further assessing the validity of the SWA instrument in both children and older adults (using DLW as the criterion).

CHAPTER 2. EXTENDED REVIEW OF LITERATURE: OBJECTIVE ACTIVITY MONITORING UNDER FREE-LIVING CONDITIONS.

The importance of physical activity for good health has been well established (US Department of Public Health, 1996). A number of professional and public health groups have published guidelines designed to facilitate awareness and promotion of PA in youth (Strong et al. 2005), adults (Haskell et al. 2007), and older adults (Nelson et al. 2007). The recent publication of the official U.S. Physical Activity Guidelines provides the most visible and definitive guidelines for physical activity (US Department of Human Health 2008). The importance of having valid, reliable and feasible tools to assess physical activity under free-living conditions is critical for evaluating compliance with guidelines and for advancing research on physical activity. While a number of assessment tools are available, accelerometry-based activity monitors have emerged as the most practical way to assess activity under free-living conditions.

This literature review will summarize current knowledge about objective activity monitoring and provide a foundation for the proposed dissertation research. The first section will review advantages and disadvantages of the most commonly used accelerometers for assessing PA under free-living conditions. The second section will describe the potential of newly developed pattern-recognition monitors for improving the accuracy of field based monitoring. This second section will mainly focus in the SenseWear Armband (SWA), the most prominent pattern-recognition monitor available (*Personal communication*, 2010).

Accelerometry-based activity monitors

Accelerometry-based activity monitors are the most commonly used method to assess PA under free-living conditions (Troiano 2005). They offer a reasonable compromise between feasibility and validity. Accelerometry-based activity monitors can be used to objectively assess frequency, duration and intensity of PA. The main limitation of accelerometers is that they are unable to capture all forms of activity (Welk et al. 2000), such as detecting energy cost increases due to carrying loads, moving up grades or strength-related efforts, or low-intensity lifestyle activities that account for the majority of the day (Bassett et al. 2000; Crouter et al. 2006; Hendelman et al. 2000).

Accelerometry-based activity monitors also present other measurement challenges. Most devices produce output in the form of activity counts (i.e., counts/min), units that are difficult to convert into usable outcome measures. Calibration equations have been developed to convert counts into more usable units (e.g., VO_2 or EE). However, studies have demonstrated limitations in cutpoints when used to assess normal free-living activities. For example, Leenders et al. (2001) evaluated the validity of three different accelerometry-based monitors (CSA, Tritrac and Yamax) to estimate PAEE under free-living conditions in a sample of women (mean age= 25.8 ± 1.6). The DLW method was employed as the criterion measure and the participants wore the monitors concurrently during 7 consecutive days. The results showed that in comparison with the DLW method, the monitors underestimated PAEE by 59%, 35% and 59% for the CSA, Tritrac and Yamax monitors, respectively.

Strath et al. (2003) assessed the accuracy of five published accelerometer regression equations to predict time spent in different intensities during free-living activities. The results showed that for light/resting intensity activities, the variation in estimation ranged between 29% underestimation to 14% overestimation of time spent in this intensity. Furthermore, during estimation of time spent in moderate intensity activities, the variation in estimation ranged between an underestimation of 60% to an overestimation of 120%. The variability is due to different protocols and activities used in the various calibration studies. Calibration equations based on locomotor movements are accurate for assessing walking and jogging but tend to underestimate the EE cost of common lifestyle tasks. Equations based on lifestyle activities, in contrast, tend to overestimate locomotor activities due to the “calibration” required to accurately assess the energy costs of lifestyle tasks. A study by Treuth et al. (2004) demonstrated that the relationship between accelerometer counts and EE varies depending on the activity being assessed. Efforts to create a single equation that fits all data points led to poor predictive accuracy so they opted to use an equation based on locomotor activities. Matthews (2005) provided a detailed review of different calibration equations and highlighted the inherent challenge of trying to calibrate these monitors using a single prediction equation.

Another challenge for accelerometry-based monitors is that it has proven difficult to determine actual wear time and monitor compliance (Catellier et al. 2005; Trost et al. 2005; Ward et al. 2005). In order to obtain accurate estimates of daily activity and EE, it is important to ensure that participants are actually wearing the monitor. However, periods of non-wear time cannot be easily distinguished from time at rest. Researchers

have developed a number of different strategies to detect and track wear time (i.e., number of consecutive minutes without activity counts). A number of studies have evaluated different assumptions (e.g., length of day, # of days, and criteria for detecting non wear time) but it has proven difficult to determine the most effective approach. Sirard and Slater (2008) evaluated different strategies for improving compliance while wearing a physical activity accelerometer (Actigraph®) in a sample of high school students. The different strategies employed included calling the participants, having the participants complete an activity log and compensating the participants with an amount contingent on number of complete (> 10 hours) days of data. The results showed that in the contingent group 96% of the participant had at least 4 out of 7 days of complete data and that these participant's compliance was significantly higher ($p = 0.04$) than in the journal (85%), phone (72%), and control (70%) participants. Furthermore, sensitivity analysis indicated that using non-zero cutpoints of 20-, 30-, 45- and 60-min criteria resulted in usable data (>4 days) for 82.6%, 83.7%, 85.9%, and 85.9% of the participants, respectively. This study demonstrates that data reduction decisions can influence reported compliance as well as summarized physical activity results. Unfortunately, there is little consensus and the problems remain (Mâsse et al. 2005; NIH Physical Activity Assessment Conference, 2009).

Despite these limitations, accelerometer-based activity monitors have demonstrated good utility for assessing overall levels of physical activity. They have limited utility for assessing individual levels of activity but are reasonably effective for assessing activity levels for group based comparisons. They have been widely used as

outcome measures for smaller studies (Sirard et al. 2008) and as criterion measures in studies validating other measures such as self-report instruments (Anderson et al. 2005). They have also been used in large surveillance research. For example, the National Health and Nutritional Examination Survey (NHANES), a large cross-sectional study of a complex multistage probability sample of US civilian, incorporated accelerometers for objective assessment of PA (Troiano et al. 2008).

In addition, accelerometers can serve as complimentary measures to allow for triangulation of outcomes. For example, Starling et al. (1999) assessed the accuracy of two self-report instruments (Minnesota Leisure Time Physical Activity Questionnaire and the Yale Physical Activity Survey) and an accelerometer (Caltrac uniaxial accelerometer) in a sample of older adults (mean age=67±9 years) using doubly labeled water (DLW) as the criterion measures. The results of the study showed different levels of accuracy and precision provided by each instrument.

Pattern-recognition monitors

The limitations of accelerometry-based activity monitors have sparked interest in alternative methodologies. A variety of approaches have been tried but most can be characterized as “pattern-recognition approaches” since they are designed to detect underlying patterns of physical activity to improve assessment. A variety of sophisticated analytical techniques have been developed, such as artificial neural networking and Hidden Markov Modeling (Bonomi et al. 2009; Bonomi et al. 2009b; Pober et al. 2006; Staudenmayer et al. 2009). These have been shown to improve accuracy of estimates but

are still limited by the use of a single sensor (acceleration). Recent advances in technology have led to the development of multi-channel devices that utilize pattern recognition algorithms to estimate PAEE and TEE. Similar to speech recognition technology, real-time monitoring of combined PA-related variables provides researchers with more detailed data on postural changes, movement and time spent in activities of varying intensities. Patterns are detected from synchronous recording and integration of various combinations of physical, biological and physiological variables from multiple sensors, which allows for more customized and specific prediction algorithms to be applied. The enhanced capabilities of pattern recognition monitors may help overcome the inherent limitations associated with traditional accelerometry-based activity monitors. Corder et al. (2007) demonstrated that pattern-recognition monitors provide more accurate estimates of PA than commonly used accelerometry-based activity monitors. Furthermore, pattern recognition monitors provide an opportunity to enhance data collection of free-living activities with increased precision and detail while facilitating the process for both researchers (i.e., user friendly software and data processing) and study participants (i.e., less cumbersome than DLW method) .

Previous research studies have assessed the validity of the Intelligent Device for Energy Expenditure and Activity (IDEEA), one of the first pattern-recognition activity monitors. The IDEEA monitor is a single-unit, portable system comprised of 5 accelerometer-based sensors (worn on the chest, thighs and sole of the feet) and a microprocessor/data storage unit that detects postural changes and performed activities (i.e., sitting, standing, and walking or running at differing speeds). Two studies conducted

on normal and overweight adults performing a range of sedentary and active tasks compared PAEE estimates between IDEEA and indirect calorimetry (Zhang et al. 2004; Rothney et al. 2007). The results of both studies support the accuracy of IDEEA's algorithms for detection of specific activities and resulting PAEE estimates. However, the high financial costs and cumbersomeness of the IDEEA monitor are a clear limitation for the assessment of PA in large samples under free-living conditions.

The Actiheart monitor (CamNtech, Cambridge, UK) simultaneously combines heart rate recordings and accelerometry (HR+M) to improve the precision of quantitative information regarding the nature and patterns of daily PAs in adults (Brage 2005; Strath 2005). The Actiheart is worn on the chest via electrodes placed left of the sternum and a second one placed parallel, on the mid-clavicular line at the third intercostal space. The device utilizes branched-equation modeling to determine the most accurate measure of PAs estimated by HR, accelerometry or the combination of the two. Several studies reported good validity while evaluating the Actiheart monitor under laboratory conditions (Brage et al 2005; Corder et al. 2005; Thompson et al. 2006). A previous study supported the validity of the Actiheart monitor for sedentary, household and leisure-time activities (Crouter et al. 2008) in a sample of adults. The protocol was carried under laboratory conditions and the Actiheart's PAEE estimates were within 0.09 kJ/kg/min (mean error=0.02 kJ/kg/min, 95% prediction interval=-0.17, 0.22 kJ/kg/min), compared to indirect calorimetry. The limitations associated with the device are directly related to the limitation observed in HR monitors, showing a tendency to lose HR values due to outside electronic interference caused by electromagnetic radiation, poor electrode adhesion, or

physical artifacts. To our knowledge, no previous study has assessed the validity of the Actiheart under free-living conditions.

SenseWear Pro Armband (SWA) (BodyMedia®, Pittsburgh, PA. www.bodymedia.com)

The SWA is a wireless, multi-sensor monitor that integrates information from a 2-axis accelerometer (the newer version of the monitor includes a 3-axis accelerometer) with a variety of heat-related sensors (i.e., heat flux, skin temperature, near-body ambient temperature, and galvanic skin response). The monitor weighs 79 g and has the memory capacity to record up to 14 days of continuous data (sampling by min.) or 2-hrs if sampling by seconds. The software allows for the following output variables: TEE, PAEE, PA duration and intensity (METs), sleep duration, and step counts. The device is worn around the right upper arm, positioned over the triceps muscle, midway between the acromion and olecranon processes. The integration of heat-related sensors allows the SWA to estimate the energy cost of complex and upper body PAs that may not be detected by hip-worn accelerometers. For example, the added work required to carry objects or walk up a grade can be detected and estimated from the increased heat production. The galvanic skin response (GSR) also provides unique information about activity and makes it possible to differentiate between periods of sleep. The GSR reflects changes in the skin's electrical properties (measured in units of ohms) due to sweat gland activity and psychological stimulus (e.g., upon awakening from sleep state) measured from two points on the participant's arm. Higher activity/stimulus instantaneously (0.2-0.5 sec) increases skin conductance, changing the balance of positive and negative ions in

the secreted fluid, which decreases skin resistance. Reduced activity/stimulus has the opposite effect.

The SenseWear software calculates EE for each minute of data using a Naive Bays classifier matching the sensor data to the activity class that best describes the current minute. The different activity classes include: walking, running, stationary bike, road bike, rest, resistance exercise, and other activity. Each activity class is linked to a linear regression model mapping the sensor values and body parameters to EE. Separate regression models are utilized for subjects 18 years of age or younger, and for those older than 18 years. Kilocalories and metabolic equivalents (METs) are converted using the equation $\text{METs} = \text{kcal} / \text{hour} / \text{kg}$. The inputs to the Naive Bays classifier and the regression models include the data recorded in the armband and derived inputs such as the standard deviation of the data over a number of minutes before and after the minute in question (*Personal communication*, June 2009).

A major advantage of the SWA is the minimal burden on researchers and participants. The device is attached to the upper arm with an elastic Velcro strap (available in three sizes), making it easy and comfortable to wear. Because the monitor is attached directly to the skin, the SWA also automatically detects and records wear time, and provides tools for interpreting gaps in data that may occur when showering or swimming. This is a major advancement over traditional accelerometers since it addresses the issues with compliance and monitoring time that plague traditional accelerometry-based devices. The time-stamp feature allows participants to ‘mark’ start

and stop times of specific PAs or events, providing researchers the opportunity to easily select data intervals of particular interest.

Another unique aspect of the SWA is that the proprietary algorithms are continuously being improved by the manufacturer in an effort to reduce error. Since the first study using this device was published in 2004 (Fruin et al. 2004), the software (and PAEE prediction algorithms) has evolved from version 1.0 to version 6.1. The release of new software has made it difficult to compare results across studies, but it is reasonable to assume that the latest software versions maintain the positive characteristics observed in preceding versions, while incorporating additional capabilities to the device. Past studies have documented improvements in the accuracy of the EE estimations following the release of new algorithms (Jakicic et al. 2004; Calabro et al. 2009b) but the use of proprietary algorithms has also prevented researchers from understanding and studying the integration of data in a more direct way. A detailed review of previous work with the SWA is provided below.

Validation of the SenseWear Pro Armband (SWA)

An increasing number of studies have evaluated the validity of the SWA under a variety of conditions and settings. These trends reveal the increased interest and acceptance of the SWA by PA measurement researchers.

Previous studies have validated the SWA in a variety of age groups (children, young, middle-aged, and older adults), as well as in samples with adverse health conditions [i.e., morbidly obese (mean BMI > 40.0 kg/m²), type 2 diabetics, cancer

patients and cystic fibrosis patients]. Results from these studies show high correlations between the SWA and indirect calorimetry (Cereda et al. 2007; Malavolti et al. 2007; Papazoglou et al. 2006; Dwyer et al. 2009; Calabro et al. 2009; Berntsen et al. 2008) and DLW (Mignault et al. 2005; St-Onge et al. 2007; Arvidsson et al. 2009b), although results vary for specific activities and different populations.

An early study by Jakicic et al. (2004) evaluated the accuracy of the Armband's PAEE estimates in 40 normal-weight, young adults, at rest and during different modes of activity. Preliminary results were not impressive, but the authors found good accuracy when data were used to train the algorithms to detect the underlying patterns in the data. The validation results showed non-significant underestimations for treadmill walking ($-2.8 \pm 9.4\%$), stepping ($-0.9 \pm 11.9\%$), and arm ($-3.8 \pm 9.9\%$) and cycle ergometry ($-0.9 \pm 10.7\%$).

Two studies performed on lean, young adults assessed the accuracy of PAEE estimates during cycle ergometry (Fruin et al. 2004) and treadmill walking and running (King et al. 2004). Fruin et al. (2004) reported no significant differences in mean PAEE estimates for 13 participants during 40 minutes of cycle ergometry. However, the SWA was poorly correlated with the indirect calorimetry data ($r = 0.03-0.12$), and large errors were observed in individuals with the highest and lowest PAEE. King et al. (2004) compared the SWA against indirect calorimetry and 4 accelerometers (CSA, RT3, TriTrac-R3D and BioTrainer), and the SWA showed moderate to high correlations with IC during various speeds of treadmill walking and running in 21 participants ($r = 0.50-0.84$). However, the PAEE estimates were significantly greater than the indirect

calorimetry ($p < 0.001$). Nevertheless, researchers concluded that the SWA was the best monitor estimating total EE at most speeds (except for slow walking).

A recent study by Berntsen et al. (2010) assessed the validity of the SWA during 2 hours of free-living activity using a portable metabolic analyzer as the criterion measure. Researchers reported underestimations for TEE (-9%). Furthermore, the SWA overestimated ($p=0.02$) and underestimated ($p < 0.001$) moderate intensity PA (≥ 3 METS and < 6 METS) and very vigorous intensity PA (> 9 METS).

Two other studies compared TEE estimates between the SWA (software version 4.02 in both studies) and DLW over a 10-day period (Mignault et al. 2005; St-Onge et al. 2007). Mignault et al. (2005) assessed TEE in 6 older adults (mean age = 56.5 ± 6.0 years) with type II diabetes. The authors reported high correlations ($r = 0.97$; $p < 0.0001$) and non-significant differences in PAEE estimates. St-Onge et al. (2007) compared TEE and PAEE estimates in 45 healthy men and women with a wider age range (ages 20-78 yrs). The SWA significantly underestimated TEE and PAEE (-117 and -225 kcal/day, respectively, $p < 0.01$). Compared to DLW, overall agreement for estimates of TEE ($R^2 = 0.74$, $SEE = 189$ kcal/day) was greater than estimates of PAEE ($R^2 = 0.49$, $SEE = 179$ kcal/day), but the values were significantly different for both ($p < 0.01$).

Johannsen et al. (2010) assessed the accuracy of the SWA (including the latest version of the software, version 6.1), and the newest version of the monitor (SWA Mini, SenseWear Software 7.0) against criterion data from the doubly-labeled water (DLW) technique in a sample of 30 adults (mean age = 38.2 ± 10.6). Absolute error rates of $\sim 8.0\%$ were observed for the two versions of the SWA. In addition, mean TEE estimates

differences were 112 kcal/ day and 22 kcal/ day for the SWA and the SWA Mini, respectively.

Some studies have reported differences in accuracy for specific segments of the population. Papazoglou et al. (2006) assessed 142 obese individuals (mean BMI = 42.3 ± 7.0 kg/m²). Although resting energy expenditure (REE) estimates from the SWA and indirect calorimetry were highly correlated ($r = 0.88$, $p < 0.001$), the SWA underestimated REE by a mean of 8.8% and overestimated PAEE during cycle ergometry (19.0%), stair stepping (30.6%) and treadmill walking (31.4%). Dwyer et al. (2009) evaluated the SWA during level and graded walking (treadmill) in 17 cystic fibrosis patients. Researchers reported high correlations with indirect calorimetry for level ($r = 0.89$; $p < 0.001$) and graded walking ($r = 0.87$; $p < 0.001$), however, the SWA significantly overestimated at low intensities and underestimated at higher intensities ($p < 0.001$).

Two studies reported good validity for the Armband's REE estimates (Cereda et al. 2007; Malavolti et al. 2007). Cereda et al. (2007) assessed REE in 10 cancer patients (mean age = 56.6 ± 13.3 yrs) while Malavolti et al. (2007) evaluated REE in 99 normal-weight adults (mean age = 38 ± 14 yrs). Both studies reported high correlations ($r = 0.84$ - 0.86 , $p < 0.0001$) and non-significant differences between REE estimates and values from indirect calorimetry. Because REE has major effects on the estimates of TEE, additional research is needed with other populations to clarify the accuracy of the REE estimates in different populations.

Studies have also reported discrepant findings when the SWA has been used in children. Arvidsson et al. (2007) assessed the validity of the SWA's REE and PAEE estimates in 20 boys and girls (ages 11-13 yrs) performing a wide range of activities. Significant underestimates were reported for most activities (rest, playing games on a mobile phone, stepping, cycle ergometry, jumping on a trampoline, basketball, and treadmill walking and running). In a later study by the same laboratory (Arvidsson et al. 2009a), researchers reported similar underestimations in PAEE during resting, sitting, stationary cycling, basketball, stair walking, and running. The researchers observed that the underestimates in PAEE increased with increasing intensity in the activities. In contrast, Dorminy et al. (2008) reported consistent overestimation of EE for a variety of tasks. The manufacturer has recently modified the prediction algorithms for youth based on new data and a previous study suggests improved accuracy (Calabro, et al. 2009b). The existing algorithms were found to overestimate EE by 32% but the average error with the newly developed algorithm was only 1.7%. In a recent study by Arvidsson et al. (2009b), assessing TEE with DLW in children over a 14-day period, the new algorithms (software version 6.1) showed a clear improvement over the previously developed algorithms (version 5.1). The new algorithms developed specifically for youth appear to have improved utility and accuracy although additional testing is required.

Two studies, from the same research group, have assessed the validity of the SWA using DLW as the gold standard over a 10-day period (Mignault et al. 2005; St-Onge et al. 2007). In the first study, type 2 diabetic patients wore the device over the testing period under free-living conditions. The researchers concluded that the SWA is an

acceptable device to measure total daily EE accurately (Mignault et al. 2005). The findings from the second study (St-Onge et al. 2007) supported the reasonable concordance of the SWA with DLW's daily EE estimates for adults under free-living conditions. Cereda et al (2007), as previously mentioned, tested the validity of the monitor for obtaining REE measurements in cancer patients. In addition, Cereda and colleagues (2007) compared total daily EE obtained from the SWA and indirect calorimetry, reporting no apparent bias between overall values from the estimates and a significant correlation between them ($r=0.68$; $p=0.001$). In the conclusions, the researchers suggested the usefulness of the SWA to estimate total daily EE in cancer patients. Welk et al. (2007) compared the SWA with another pattern recognition monitor (IDEEA) in college age participants ($N=30$, mean age: 24.9 ± 6.1 years, BMI: 25.9 ± 5.6 kg/m²), during their daily-living activities. Researchers reported good agreement between the two monitors under a wide variety of activities.

In a recent study in our laboratory, we compared the validity of different activity monitors for EE estimation of light intensity activity in young adults. This is an important issue since recent research has demonstrated that time spent in sedentary activities may have independent effects on health (Martinez-Gomez et al. 2009, Warren et al. 2010). Studies have also demonstrated health benefits associated with the accumulation of light intensity activity (Healy et al. 2007; Healy et al. 2008; Levine et al 1999). We evaluated the relative ability of various PA assessment tools for capturing lower intensity activities. Data from 3 pattern-recognition monitors (SWA, v. 6.1; Mini, v. 7.0 and Actiheart monitor), and 2 accelerometry-based activity monitors (ActiGraph GT3X and ActivPAL)

were compared to a portable metabolic analyzer (Oxycon Mobile) during 60 minutes of free-living activity. The results showed that the Mini and SWA monitors provided more accurate estimates of EE during light to moderate intensity free-living activities compared to other activity monitors (*Personal communication*, April 2011).

Collectively, the above mentioned studies support the validity of the SWA as a method to assess TDEE under free-living conditions. From the available literature, it appears that the discrepancies between the EE estimates from the SWA and the different comparison methods are more evident in unique, homogenous populations rather than in healthy adults. Papazoglou et al. (2006) reported that the SWA underestimated REE and highly overestimated EE during exercise sessions including cycle ergometry, stair stepping and treadmill walking in obese individuals. Furthermore, in a study by Arvidsson et al. (2007) in children (~12 years), researchers reported significant underestimation of PAEE for most of the activities (Rest, playing games on mobile phone, stepping, cycle ergometry, jumping on trampoline, playing basketball, and walking and running on a treadmill). It is possible that the SWA proprietary algorithms might have been developed using data obtained with lean adults, and consequently, estimates from the SWA in populations with other characteristics will be less accurate for most activities. Additionally, the REE estimation from the SWA might differ from the true REE values and, therefore, partially explain some of the error observed. Therefore, it seems evident that in order to be able to assess populations with different anthropometric characteristics, specific algorithms should be developed and included in the SWA software.

The available literature on the SWA supports the validity of the instrument to assess free-living activity in healthy adults. Additional work on the SWA should concentrate in the development of specific algorithms to accurately estimate EE in different populations (i.e.: Overweight/obese, children, older adults). In addition, improvements in the algorithms used to estimate EE during certain activities should be also considered.

References

Adams SA, Matthews CE, Ebbeling CB, et al. The effect of social desirability and social approval on self-reports of physical activity. *Am J Epidemiol* 2005; 161:389–98.

Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, O'Brien WL, Bassett DR, Jr., Schmitz KH, Emplaincourt PO, Jacobs DR, Jr., and Leon AS.

Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc.* 2000; 32: S498-504.

Anderson CB, Hagströmer M, Yngve A. Validation of the PDPAR as an adolescent diary: effect of accelerometer cut points. *Med Sci Sports Exerc* 2005;37(7):1224-30.

Arvidsson D, Slinde F, Larsson S, Hulthen L. Energy cost of physical activities in children: validation of SenseWear Armband. *Med Sci Sports Exerc.* 2007;39(11):2076-2084.

a-Arvidsson D, Slinde F, Larsson S, Hulthén L. Energy cost in children assessed by multisensor activity monitors. *Med Sci Sports Exerc.* 2009;41(3):603-11.

b-Arvidsson D, Slinde F, Hulthen L. Free-living energy expenditure in children using multi-sensor activity monitors. *Clinical Nutr.*2009; 28:305-312.

Bassett, DR Jr., Ainsworth BE, Swartz AM, Strath SJ, O'Brien WL, King GA. Validity of four motion sensors in measuring moderate intensity physical activity. *Med Sci Sports Exerc* 2000;32(9 Suppl):S471-480.

Berntsen S, Hageberg R, Aandstad A, Mowinckel P, Anderssen SA, Carlsen KH, Andersen LB. Validity of physical activity monitors in adults participating in free living activities. *Br J Sports Med*. 2010; 44(9):657-64.

Bland JM and Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 1: 307-310.

a-Bonomi AG, Goris AH, Yin B, Westerterp KR. Detection of type, duration, and intensity of physical activity using an accelerometer. *J Appl Physiol*. 2009; 107: 655-661.

b-Bonomi AG, Plasqui G, Goris AH, Westerterp KR. Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer. *Med Sci Sports Exerc*. 2009; 41: 1770-1777.

Brage S, Brage N, Franks PW, Ekelund U, and Wareham NJ. Reliability and validity of the combined heart rate and movement sensor Actiheart. *Eur J Clin Nutr.* 2005; 59: 561-70.

a- Calabro MA, Welk GJ, Carriquiry AL, Nusser SM, Beyler NK, Mathews CE. Validation of a computerized 24-hour physical activity recall (24PAR) instrument with pattern-recognition activity monitors. *J Phys Act Health.* 2009; 6(2):211-20.

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

Calabro MA, Lee J, Harrington, D, Saint-Maurice, PF, Yoo, H, and Welk, GJ. Validity of Physical Activity Monitors for Assessing Lower Intensity Activity in Adults. 2011 (*Manuscript in preparation*).

Catellier DJ, Hannan PJ, Murray DM, Addy CL, Conway TL, Yang S, Rice JC. Imputation of missing data when measuring physical activity by accelerometry. *Med Sci Sports Exerc.* 2005; 37(11 Suppl): S555-S562.

Cereda E, Turrini M, Ciapanna D, Marbello L, Pietrobelli A, Corradi E. Assessing energy expenditure in cancer patients: A pilot validation of a new wearable device. *J Parenter Enteral Nutr* 2007; 31(6):502-507.

Corder K, Brage S, Mattocks C, et al. Comparison of two methods to assess PAEE during six activities in children. *Med Sci Sports Exerc.* 2007; 39(12): 2180–8.

Corder, K., Brage, S., Wareham, N.J., and U. Ekelund. Comparison of PAEE from combined and separate heart rate and movement models in children. *Med Sci Sports Exerc* 2005; 37(10): 1761-7.

Crouter SE, Churilla JR, Bassett DR Jr. Accuracy of the Actiheart for the assessment of energy expenditure in adults. *Eur J Clin Nutr.* 2008; 62: 704-11.

Crouter SE, Churilla JR, Bassett DR Jr. Estimating energy expenditure using accelerometers. *Eur J Appl Physiol.* 2006; 98(6): 601-612.

Dorminy CA, Choi L, Akohoue SA, Chen KY, Buchowski MS. Validity of a multisensor armband in estimating 24-h energy expenditure in children. *Med Sci Sports Exerc.* 2008; 40(4):699-706.

Dwyer TJ, Alison JA, McKeough ZJ, Elkins MR, Bye PT. Evaluation of the SenseWear activity monitor during exercise in cystic fibrosis and in health. *Respir Med.* 2009; 21.

Fruin ML, Rankin JW. Validity of a multi-sensor armband in estimating rest and exercise energy expenditure. *Med Sci Sports Exerc.* 2004; 36(6): 1063-1069.

Haskell WL et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Med Sci Sports Exerc.* 2007; 39(8): 1423-34.

Healy GN, Dunstan DW, Salmon J, Cerin E, Shaw JE, Zimmet PZ, Owen N. Objectively measured light-intensity physical activity is independently associated with 2-h plasma glucose. *Diabetes Care.* 2007; 30(6): 1384-9.

Healy GN, Dunstan DW, Salmon J, Cerin E, Shaw JE, Zimmet PZ, Owen N. Breaks in sedentary time: beneficial associations with metabolic risk. *Diabetes Care.* 2008; 31(4): 661-6.

Hendelman D, Miller K, Baggett C, Debold E, Freedson P. Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Med Sci Sports Exerc.* 2000; 32(9 Suppl): S442-449.

Jakicic JM, Marcus M, Gallagher KI, Randall C, Thomas E, Goss FL, Robertson RJ. Evaluation of the SenseWear Pro Armband to assess energy expenditure during exercise. *Med Sci Sports Exerc.* 2004; 36(5): 897-904.

Johannsen DL, Calabro MA, Stewart J, Franke W, Rood JC, Welk GJ. Accuracy of armband monitors for measuring daily energy expenditure in healthy adults. *Med Sci Sports Exerc.* 2010; 42(11): 2134-40.

Jones RH, Sonko BJ, Miller LV, Thureen PJ, Fennessey PJ. Estimation of doubly labeled water energy expenditure with confidence intervals. *Am J Physiol Endocrinol Metab.* 2000; 278: E383–E389.

King GA, Torres N, Potter C, Brooks TJ, Coleman KJ. Comparison of activity monitors to estimate energy cost of treadmill exercise. *Med Sci Sports Exerc.* 2004; 36(7): 1244-1251.

Leenders NY, Sherman WM, Nagaraja HN, Kien CL. Evaluation of methods to assess physical activity in free-living conditions. *Med Sci Sports Exerc.* 2001; 33(7):1233-40.

Levine JA. Nonexercise activity thermogenesis (NEAT): environment and biology. *Am J Physiol Endocrinol Metab.* 2004; 286: E675-685.

Levine JA, Eberhardt NL, Jensen MD. Role of nonexercise activity thermogenesis in resistance to fat gain in humans. *Science*; 1999; 283.

Mâsse LC, Fuemmeler BF, Anderson CB, Matthews CE, Trost SG, Catellier DJ, Treuth M. Accelerometer data reduction: a comparison of four reduction algorithms on select outcome variables. *Med Sci Sports Exerc.* 2005; 37(11): S544-54.

Matthew CE. Calibration of accelerometer output for adults. *Med Sci Sports Exerc.* 2005; 37: S512-522.

Matthews CE, Freedson PS, Hebert JR, Stanek EJ III, Merriam PA, Ockene IS. Comparing physical activity assessment methods in the seasonal variation of blood cholesterol study. *Med Sci Sports Exerc.* 2000; 32: 976–984.

Malavolti M, Pietrobelli A, Dugoni M, Poli M, Romagnoli E, De Cristofaro P, Battistini NC. A new device for measuring resting energy expenditure (REE) in healthy subjects. *Nutr Metab Cardiovasc Dis.* 2007;17(5):338-343.

Mignault D, St-Onge M, Karelis AD, Allison DB, Rabasa-Lhoret R. Evaluation of the Portable HealthWear Armband: a device to measure total daily energy expenditure in free-living type 2 diabetic individuals. *Diabetes Care.* 2005; 28(1): 225-227.

Mirwald RL, Baxter-Jones ADG, Bailey DA, Beunen GP. An assessment of maturity from anthropometric measurements. *Med Sci Sports Exerc.* 2002; 34(4); 689-694.

Montoye HJ, Kemper HC, Saris WHM, and Washburn RA. *Measuring Physical Activity and Energy Expenditure*. Champaign, IL: Human Kinetics, 1996, pp. 17-25.

Nelson MF, Rejeski WJ, Blair SN, Duncan PW, Judge JO, King AC, Macera CA, Castaneda-Sceppa C. Physical activity and public health in older adults: recommendation from the American College of Sports Medicine and the American Heart Association. *Med Sci Sports Exerc.* 2007; 39(8): 1435-45.

Papazoglou D, Augello G, Tagliaferri M, Savia G, Marzullo P, Maltezos E, Liuzzi A. Evaluation of a multisensor armband in estimating energy expenditure in obese individuals. *Obesity* (Silver Spring). 2006; 14(12): 2217-2223.

Pober DM, Staudenmayer J, Raphael C, Freedson PS. Development of novel techniques to classify physical activity mode using accelerometers. *Med Sci Sports Exerc.* 2006; 38(9): 1626-34.

Racette SB, Schoeller DA, Luke AH, Shay K, Hnilicka J, and Kushner RF. Relative dilution spaces of 2H- and 18O-labeled water in humans. *Am J Physiol.* 1994; 267: E585-590.

Ross WD, Marfell-Jones MJ. Kinanthropometry. In MacDougall JD, Wenger HA, Green HJ, eds. *Physiological Testing of the High-Performance Athlete*, pp 223-308. Champaign, Illinois: Human Kinetics Books, 1991.

Rothney MP, Neumann M, Beziat A, Chen KY. An artificial neural network model of energy expenditure using nonintegrated acceleration signals. *J Appl Physiol.* 2007; 103(4): 1419-27.

Sallis J F, Haskell W, Wood P, et al. Physical activity assessment methodology in the Five-City Project. *Am J Epidemiol.*1985; 121:91-106.

Schoeller DA. Measurement of energy expenditure in free-living humans by using doubly labeled water. *J Nutr.* 1988; 118: 1278-1289.

Sherman SE, D'Agostino RB, Cobb JL, Kannel WB. Does exercise reduce mortality rates in the elderly? Experience from the Framingham Heart Study. *Am Heart J.* 1994; 128: 965-972.

Sirard JR, Kubik MY, Fulkerson JA, Arcan C. Objectively measured physical activity in urban alternative high school students. *Med Sci Sports Exerc.* 2008; 40(12): 2088-95.

Sirard, JR, Slater, ME. Compliance with wearing physical activity accelerometers in high school students. *Med Sci Sports Exerc.* 2008; 40 (5): S208.

Slootmaker SM, Schuit AJ, Chinapaw MJM, Seidell JC, van Mechelen W. Disagreement in physical activity assessed by accelerometer and self-report in subgroups of age, gender, education and weight status. *Int J Behav Nutr Phys Act.* 2009; 6: 17.

Staudenmayer J, Pober D, Crouter SE, Bassett DR, and Freedson P. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. *J Appl Physiol.* 2009; 107(4): 1300-7.

Starling RD. Use of doubly labeled water and indirect calorimetry to assess physical activity. In: Welk GJ. *Physical Activity Assessments for Health-Related Research.* Champaign, Ill: Human Kinetics; 2002: 197-209.

Starling RD, Matthews DE, Ades PA, Poehlman ET. Assessment of physical activity in older individuals: a doubly labeled water study. *J Appl Physiol.* 1999; 86(6): 2090–2096.

St-Onge M, Mignault D, Allison DB, Rabasa-Lhoret R. Evaluation of a portable device to measure daily energy expenditure in free-living adults. *Am J Clin Nutr.* 2007; 85(3): 742-749.

Strath SJ, Bassett DR Jr, Swartz AM. Comparison of MTI accelerometer cut-points for predicting time spent in physical activity. *Int J Sports Med.* 2003; 24(4): 298-303.

Strath SJ, Brage S, Ekelund U. Integration of physiological and accelerometer data to improve physical activity assessment. *Med Sci Sports Exerc.* 2005; 37(11): S563–S571.

Strong WB et al. Evidence based physical activity for school-age youth. *J Pediatr.* 2005; 146: 732-7.

Treuth MS, Schmitz K, Catellier DJ, McMurray RG, Murray DM, Almeida MJ, Going S, Norman JE, and Pate RJ E. Defining accelerometer thresholds for activity intensities in adolescent girls. *Med Sci Sports Exerc.* 2004; 36 (7): 1259-1266.

Troiano RP. A timely meeting: objective measurement of physical activity. *Med Sci Sports Exerc.* 2005; 37(11): S487–9.

Troiano RP et al. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc.* 2008; 40: 181–8.

Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. *Med Sci Sports Exerc.* 2005; 37(11 Suppl): S531-543.

Thompson D, Batterham AM, Bock S, Robson C, Stokes K. Assessment of low-to-moderate intensity physical activity thermogenesis in young adults using synchronized heart rate and accelerometry with branched-equation modeling. *J Nutr.* 2006; 136: 1037-42.

US Department of Health and Human Services. Physical activity and health: A report of the Surgeon General. 1996.

US Department of Health and Human Services. 2008 Physical activity guidelines for Americans. Washington (DC): US Department of Health and Human Services; 2008. Available from: United States Department of Health and Human Services.

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 Suppl): S582-S588.

Washburn RA, Jacobsen DJ, Sonko BJ, Hill JO, Donnelly JE. The validity of the Stanford seven-day physical activity recall in young adults. *Med Sci Sports Exerc.* 2003; 35 (8):1374–1380.

Welk GJ, Blair SN, Wood K, Jones S, & Thompson KW. . A comparative evaluation of three accelerometry-based physical activity monitors. *Med Sci Sports Exerc.* 2000; 32(9), S489-S497.

Welk GJ, McClain JJ, Eisenmann JC, Wickel EE. Field validation of the MTI Actigraph and BodyMedia armband monitor using the IDEEA monitor. *Obesity (Silver Spring).* 2007; 15(4): 918-928.

Westerterp KR, Brouns F, Saris WH, and ten Hoor F. Comparison of doubly labeled water with respirometry at low- and high-activity levels. *J Appl Physiol.* 1988; 65: 53-56.

Zhang K, Pi-Sunyer FX, Boozer CN. Improving energy expenditure estimation for physical activity. *Med Sci Sports Exerc.* 2004; 36(5): 883-9.

CHAPTER 3. PHYSICAL ACTIVITY ASSESSMENT IN OLDER ADULTS UNDER FREE-LIVING CONDITIONS.

A manuscript to be submitted for publication in the *Journal of Gerontology*

M. A. Calabró, J.M. Stewart, W.D. Franke, G. J. Welk.

Abstract

The importance of physical activity (PA) in maintaining overall health with advancing age has been well documented. Older adults are less likely to perform structured bouts of PA, therefore, there is a need to develop physical activity measurement instruments that can capture patterns of activity more typical for older adults. **Purpose:** The purpose of the study was to evaluate the validity of two SenseWear Armband monitors (The Pro₃ (SWA) and the recently released Mini (Mini)) and a self-report instrument under free-living conditions in older adults. **Methods:** Participants in the study (20 healthy adults aged 60-78yrs) wore both monitors for 14 consecutive days, including sleeping time. Estimates of total energy expenditure (TEE) from the SenseWear monitors were computed using the latest algorithms (version 2.2 algorithms, available in the SenseWear software version 6.1 and 7.0). The estimates were compared to estimates derived from the doubly labeled water methodology (DLW), and a self-report instrument (7D-PAR), using standard measurement agreement procedures. **Results:** Total EE comparisons showed non-significant differences for the Mini (11.6 (0.4%) kcal·day⁻¹), the SWA (77.7 (2.9%) kcal·day⁻¹) and the 7D-PAR (-134.9 (5%) kcal·day⁻¹) compared to the DLW method. The absolute error rates (computed as average absolute value of the individual errors) were

very similar for the Mini (8.4%) and for the SWA (6.5%) but higher for the 7D-recall (12.8%). Pearson product-moment correlations were high for the activity monitors (Mini: $r=0.85$, SWA: $r=0.91$) and slightly lower for the 7D-PAR ($r=0.75$). Bland-Altman plots for TEE did not show systematic bias for any of the measurements. Activity EE absolute error rates for the monitors were higher (Mini=21.9% and SWA=17.1%) but much larger for the self-report instrument (7D-PAR=33.5%). The 7D-PAR showed systematic bias trends ($r=0.57$) in the Bland-Altman plot for AEE. **Conclusions:** The SenseWear Pro₃ and SenseWear Mini Armband monitors were found to provide reasonably valid estimates of EE and PAEE in older adults, under free-living conditions. The 7D-PAR provided reasonable group level estimates but less accurate estimates of EE for individual estimation.

Key words: Energy expenditure, physical activity, activity monitor.

Introduction

The importance of physical activity (PA) in maintaining overall health with advancing age has been well documented (1). Most evidence has been obtained from data using self-report instruments. These tools have been shown to be prone to bias and error (2) so there is a need to develop more effective physical activity measurement instruments. The use of accelerometry-based activity monitors has become an accepted practice but they are more effective for capturing locomotor activities. Studies have demonstrated poor validity for assessing lower intensity, lifestyle activities (3). Older adults are less likely to perform structured bouts of physical activity (4) so alternative monitoring techniques may be needed to assess physical activity in this population.

Pattern recognition monitors, that combine information from multiple sensors and have similar cost to accelerometers, offer promise for assessing the unique activity patterns in older adults. The combination of multiple sensors allows for the detection of PA patterns and the application of activity specific algorithms to provide accurate estimation of energy expenditure (EE). Pattern-recognition monitors have been shown to provide more accurate estimates of PA than commonly used accelerometers (5, 6).

The SenseWear Pro₃ Armband (SWA, *Bodymedia Inc.*, Pittsburgh, PA) is a pattern-recognition monitor that integrates information from a 2-axis accelerometer and a variety of additional sensors (i.e., heat flux, skin temperature, near-body ambient temperature, and galvanic skin response). The integration of heat-related sensors with acceleration allows the monitor to estimate the energy cost of complex PAs that may not

be detected by hip-worn accelerometers. Previous laboratory validation studies have supported the validity of the SWA for assessing PA and EE in adults (7, 8, 9, 10, 11). A recent doubly-labeled water study further supported the validity of the SWA for estimating EE in healthy young and middle-aged adults (12). Estimates from the SWA were within 8% absolute error from the estimates obtained from the DLW method. Additionally, the researchers reported a high level of agreement between the methods (ICC>0.80).

To date, the validity of the most current version of the SWA software (Software V.6.1) has not been formally tested in older adults. In addition, the validity of the recently released SenseWear Mini Armband (Mini, Software V.7.0) monitor has not been tested in older adults to date. Therefore, the primary purpose of the present study was to evaluate the validity of the SWA and the Mini for measuring total daily EE in older adults under free-living conditions using direct comparisons from the DLW method. The doubly labeled water (DLW) method is widely accepted as the “gold standard” for energy expenditure (EE) assessment under free-living conditions (13, 14) so this provides an ideal measure for assessing the validity of the SWA in this population.

A secondary purpose of this study was to evaluate the validity of a commonly used self-report instrument against the temporally matched data from the SWA and DLW. The DLW provides the optimal way to validate the SWA but the SWA is a better outcome measure to evaluate the validity of the self-report measure. The reason for this is that the SWA can provide estimates of the amount of time spent in different intensities of activity and data can also be partitioned by time. The inclusion of the self-report data also

provides valuable contextual data to interpret and potentially explain error in the SWA data (as assessed relative to the DLW). Triangulation of data from the three sources will help to advance research on the assessment of physical activity and energy expenditure in this population.

Methods

Participants

A sample of 20 healthy older adult (age range: 60-80yrs) men and women was recruited to participate in the study through word of mouth. Participants were screened for health conditions that may have prevented them from being active, that would impact the function of the monitors, or affect the ability to self-report activity behavior.

Exclusion criteria included suffering from any health condition that may have prevented the participant from being active and the use of supplemental oxygen or medical devices. Additionally, individuals taking diuretics or thyroid medication were excluded from the study to avoid a confounding effect with the DLW method.

Approval from the Institutional Review Board was obtained to ensure that the protocol met established procedures. Participants were informed about the procedures and purposes of the study before signed consent was obtained.

Physical Activity Assessment methods

Doubly Labeled Water

The DLW method is a non-invasive technique that, via ingestion of isotopic tracers (deuterium and Oxygen -18) gradually eliminated from the body, enables estimation of carbon dioxide production and subsequent estimation of oxygen consumption (using a standardized Respiratory Quotient). The method is viewed as the most accepted criterion measure of TEE (15) but additional processing and information is required to estimate activity energy expenditure (AEE), the most variable component of daily EE. Activity EE can be estimated by subtracting the resting metabolic rate (RMR), and diet induced thermogenesis (~10%), from daily EE (16). While DLW is an expensive and complicated method for large scale assessments, it provides an ideal criterion measure to evaluate the accuracy of the EE and AEE estimates from other assessment tools.

SenseWear Armband Pro₃ Armband (SWA, Model 908901 PROD2) and SenseWear Mini Armband (Mini, Model MS-SW)

The SWA is a wireless multi-sensor activity monitor that integrates motion data from two orthogonal accelerometers along with several heat related sensors (heat flux, body temperature and galvanic skin response). The monitor, worn on the upper right arm over the triceps muscle (midway between the acromion and olecranon processes), is lightweight (79-grams, 85x35x20mm) and comfortable to wear. The SWA uses a replaceable AAA battery for 14 days of use and the memory allows for 10 consecutive

days of storage. In the current study, the SWA was processed using the latest proprietary algorithms available in the software (Research Software v.6.1).

A second monitor, the SenseWear Mini Armband (Mini), is a newer (and smaller, 45.4g, 55x62x13mm) version on the Armband monitor that it is worn on the left arm. The Mini operates in a similar manner as the more established SWA but includes a triaxial accelerometer instead of two individual accelerometers. In addition, the Mini has a rechargeable battery for 7 days of use, and a memory for 14 consecutive days. The Mini uses a different software package (Research Software v.7.0) than the SWA with similar proprietary algorithms.

SenseWear Software v.6.1 and v.7.0 both use the same EE algorithm architecture (algorithm v.2.2). This EE algorithm uses the different sensors (i.e., skin temperature, galvanic skin response, heat flux, and accelerometry) to provide EE estimates for each minute of data using complex pattern-recognition algorithms. Through a Naive Bays classifier, the sensors data are matched to different activity classes (i.e., walking, running, stationary bike, road bike, rest, resistance exercise or other activity) and later linked to linear regression models, mapping the sensor values and body parameters to EE.

7-day physical activity recall (7D-PAR)

The 7D-PAR instrument has been one of the most widely used physical activity instruments in the field and has been previously validated using the DLW method (17, 18). It provides data across a full 7-day period in order to capture typical activity behavior but this advantage may be offset by the lack of precision in the data for the

individual days (19). Participants must answer questions about the number of minutes spent in moderate, hard and very hard intensity activity in the morning, afternoon and evening for each day of the week. “Moderate” activity is defined as any activity that is similar to how you feel when you are walking at a normal pace. Any activity that is similar to how you feel when you are running or jogging is defined as “very hard” intensity activity. “Hard” intensity activity is defined as any activity that falls in between the “moderate” and “very hard” categories.

Data collection procedures

Table 1 provides a schematic representation of the experimental protocol. On the first day of the study (day 0), participants reported to the campus research center (Nutrition and Wellness Research Center) following a 10-hour overnight fast (no food or drink other than water) and after collecting a baseline urine sample (Baseline A). Participants then provided a second baseline urine sample upon arrival (Baseline B) and standard anthropometric measurements were collected.

Standing height was measured to the nearest 0.1 cm with the use of a wall mounted Ayrton stadiometer (Prior Lake, MN) and with the participants barefooted. Body mass was measured with participants in light clothes and barefooted on a Cardinal Detecto electronic scale (Webb City, MO) to the nearest 0.1 kg. The body mass index (BMI) was calculated as $\text{weight (kg)}/\text{height}^2 (\text{m}^2)$.

The two monitors were initialized using the participant’s personal information (age, gender, height, weight, smoking status and handedness) and adjusted to fit tightly on the participants’ arms. The SWA and Mini monitors were placed on the right and left

arms, respectively, following manufacturer recommendations. After fitting both activity monitors, a DLW dose was administered to the participant. The dose was determined based on body weight in accordance with a standardized protocol. Participants received a 1.5 ml/kg body wt. dose of a mixture of 10% enriched H_2^{18}O and 99% enriched $^2\text{H}_2\text{O}$ (Cambridge Isotopes, Cambridge, MA). After the dose was administered, 4 urine samples were collected at 1.5, 3.0, 4.5 and 6.0 hours. Liquid consumption was monitored during those 6 hours post DLW dose ingestion. After initializing the monitors, participants then were instructed to continue their normal life while wearing both activity monitors 24 hours a day, except while doing water-related activities (i.e., showering, swimming). Participants were instructed to record non-wearing periods other than self-care periods (i.e., showering, dressing).

On Day 7 and Day 14, participants reported to the lab in a fasted state and were asked to provide additional urine samples at two time points (90 minutes apart). Body weight was measured on both days to check for changes in weight over the course of the study. Immediately after, the participants completed a 7D-PAR guided by the researchers. Resting metabolic rate (RMR) data were obtained on both days (between the specimen collection) using a metabolic analyzer (Physiodyne Max-II metabolic cart; Physiodyne Instruments, Quogue, NY). Resting metabolic rate measurements were obtained in an isolated dark room, with participants lying down in a supine position without sleeping, avoiding speaking and minimizing their movement. The first 10 minutes of resting were used for acclimatization with the participants wearing the metabolic analyzer equipment. After the acclimatization period, measurements were obtained during the subsequent 15-

minute period. The RMR value was computed on a per minute basis and expressed per day to facilitate analyses. The first 10 min of data were discarded and the last 15 minutes were averaged to obtain a single RMR estimate and then extrapolated for 24-h RMR estimations. The temperature of the room was maintained at 22° C, and the calorimeter was calibrated before every measurement for pressure and gas concentrations. Replicate RMR measures were obtained in Days 7 and 14 to ensure accuracy in the measurements.

Data Processing

Processing of Doubly Labeled Water Data

The DLW procedure involved collection and processing of urine samples on Days 0, 7 and 14. The samples were labeled and coded by time to ensure accurate processing of the data. All specimens were processed by the same research technician using standardized procedures. Duplicate urine samples of approximately 12ml were stored in tubes and frozen. Samples were sent for processing by the Pennington Biomedical Research Center (Baton Rouge, LA). At Pennington, abundance of ^{18}O was measured in duplicate on a gas isotope ratio mass spectrometer (IRMS), and $^2\text{H}_2$ abundance was measured in duplicate on the same IRMS. The ^2H and ^{18}O isotope elimination rates (k_{D} and k_{O}) were calculated using linear regression following a log transformation. Total body water (N) was determined at time zero, obtained from the regression line of the H_2^{18}O isotope. The rate of CO_2 production was calculated using the equations of Schoeller (20) and later modified (21) as follows:

$$r\text{CO}_2 \text{ (moles/d)} = (\text{N}/2.078) (1.007k_{\text{O}} - 1.041k_{\text{D}}) - 0.0246r_{\text{GF}};$$

where $r\text{CO}_2$ is the rate of carbon dioxide production; N is total body water calculated from $N_{\text{O}}/1.007$ where N_{O} is the ^{18}O dilution space; k_{O} and k_{D} represent the fractional elimination rates of ^{18}O and $^2\text{H}_2$, respectively; and r_{GF} is the rate of fractionated gaseous evaporative water loss, which is estimated to be $1.05*N(1.007k_{\text{O}} - 1.041k_{\text{D}})$. Total energy expenditure (TEE) was calculated in the following manner: $\text{TEE (kcal/d)} = 22.4 r\text{CO}_2 (3.9/\text{RQ} + 1.10)$. This formula assumes a respiratory quotient (RQ) of 0.86 which is typical for a healthy, rather low fat diet. Values were expressed per day to facilitate interpretation. Additional information about the DLW protocol and processing are published elsewhere (12).

Processing of SWA Data

During visits to the laboratory on days 7 and 14 of the protocol, data from both monitors were downloaded for memory clearance and to recharge the batteries. Individual attention was given to each data file in order to control for possible gaps in the data during the monitoring period. Each non-wearing period was compared with the information provided by the participants in order to account for all gaps in the data. Later, the identified gaps were manually filled with corresponding MET values based on the Compendium of Physical Activities by Ainsworth et al. (22). For example, data gaps attributable to showering and dressing were manually filled with a corresponding MET equivalent for “self-care activities” (2.0 METs) based on the Compendium of Physical Activities (22). Further unaccounted gaps shorter than 10 minutes, which might occur due to a loose monitor strap, were filled with average EE of the 10 minutes before the defined gap strap. Planned gaps in the data that occurred in day 7 for download and battery

change/charge were filled using MET estimates of light/resting activity since participants were completing a sedentary activity during this time. After all data gaps were accounted for in each file, TEE and daily EE were calculated by summing up the amount of EE expended by the participant over the monitoring period (14 days).

Processing of Self-Report Data

The 7D-PAR was administered to the participants in days 7 and 14 of the protocol using a paper and pencil version. Participants were asked to recall the amount of time spent in sleep, moderate, hard, and very hard physical activities during weekdays and weekend days of the previous week. The average amount of time spent in light activities each day was calculated as the difference between 24 hours and the amount of time spent in sleep, moderate, hard, and very hard activities. Total daily energy expenditure (TDEE) was calculated as the average hours per day in each activity category multiplied by a previously assigned MET value (sleep=1.0, light=1.5, moderate=4.0, hard=6.0, and very hard=10), body weight (kg) and 24hrs (1MET=1.0 Kcal/kg/hour). In order to compare the 7D-PAR values with the DLW method, an average of two 7D-PAR obtained each week of the study were averaged.

Statistical analyses

Descriptive statistics were computed for the primary PA categories to describe the characteristics of the participants and their activity profiles. Data were checked for normality to ensure that the distribution of the data would not influence the results. Physical activity level (PAL) represents multiples of basal metabolic rate, an index of TEE adjusted for body weight. It is an analogous concept to METs, and it is computed as:

$PAL = TEE \cdot RMR^{-1}$. Activity energy expenditure (AEE) values were obtained by subtracting thermic effect of foods (commonly accepted as ~10% of TEE) and resting metabolic rate (RMR) from the TEE ($AEE = [(TEE \times 0.9) - RMR]$). Mean absolute error rates were computed as the average of the absolute value of the residuals divided by the actual DLW value, multiplied by 100.

The primary statistical analyses involved evaluating the relative accuracy of the SWA and Mini monitors and the self-report measure compared with the matched data from the DLW. Statistical power was estimated for associated *t*-test at the 5% level. The associated *t*-tests were used to test differences between specific instruments, in order to compare the estimates to DLW as well as to each other. Pearson product-moment correlations were computed to evaluate the associations between the various estimates of TEE and AEE. All statistical analyses were performed using IBM SPSS Statistics 19 (IBM Corporation, Somers, NY).

Bland Altman graphical procedures (23) were used to examine agreement across the range of TEE and AEE values between DLW and the alternative methods (SWA, Mini, 7D-PAR). The DLW values were plotted on the x-axis and the difference between the estimates (i.e., DLW minus SWA) were plotted on the y-axis. Residuals were correlated with the DLW values in order to assess for any systematic bias. Confidence intervals defining the limits of agreement between DLW and the alternative methods were set at 1.96 SD from the mean difference. Differences were considered significant if $p < 0.05$.

Results

Descriptive Analyses

Twenty participants (9 males, 11 females) completed the study. Descriptive statistics for the sample population are provided in Table 2. The participants were all self-described as white and the sample included a variety of body types, with 45.0% categorized as overweight ($25 \leq \text{BMI} < 30$) and 15.0% categorized as obese ($\text{BMI} > 30$) participants. Data from 1 participant had to be discarded from the analyses due to problems with the urine collection needed for the DLW analyses. Therefore, final analyses include data from 19 participants. All the participants completed the self-report instrument (7D-PAR).

The DLW data were processed to provide an indicator of the overall activity patterns of the population. Measured RMR for the sample ranged from $1031.2 \text{ kcal}\cdot\text{day}^{-1}$ to $2847.9 \text{ kcal}\cdot\text{day}^{-1}$. The associated estimates of PAL ranged from 1.1 to 2.5, reflecting a large range of activity levels among the participants. Additional details on activity patterns were obtained by processing the data from the monitors and the self-report instrument. On average, participants wore the monitors for 97.9% (± 3.8) of the time during the 14 days of monitoring. The participant's wearing percentages ranged from 87.8% to 99.9% of protocol time. The average off-body time during the 14 days of monitoring was 208.5 ± 379.8 minutes per week (range: 14 to 2411 minutes). According to estimates from the activity monitors, participants performed an average of 133.7 (range= 7.7-470.7, Mini) and 103.6 (range= 13.9-369.0, SWA) minutes of moderate activity, and 3.8 (range= 0-27.4) and 4.7 (range= 0-24.1) minutes of vigorous activity,

respectively. The 7D-PAR reported minutes were lower for moderate activity (69.8, range=0-387.9), but higher for minutes of vigorous activity (10.8, range= 0-55.7).

Distribution of moderate and vigorous activity minutes by assessment tool are shown in Figure 1.

Analyses of Total energy expenditure

The evaluation of total energy expenditure (TEE) could be influenced by differences within the sample. Analyses of covariance (ANCOVA) were used to test for the effects of gender on estimates of TEE. The results showed no significant effects for any of the comparisons with DLW; therefore, males and females were combined for all analyses. The statistical power to detect EE differences with 19 participants was ~80%.

Estimates of TEE are provided in table 3. Group mean comparisons did not yield differences between the DLW method and the Mini [diff = 11.6 kcal·day⁻¹ (0.4 %), 95% CI=-131.9 to 155.0], the SWA monitor [diff=77.7 kcal·day⁻¹ (2.9%), 95% CI=-37.2 to 192.6] and the 7D-PAR [diff=-134.9 kcal·day⁻¹ (5.0%), 95% CI=-356.2 to 86.5]. Mean absolute error values with the DLW method were 8.4%, 6.5% and 12.8%, for the Mini, SWA and 7D-PAR, respectively.

Pearson product-moment correlations between the DLW method TEE and the comparison methods were high in all cases ($r=0.85$, $r=0.91$ and $r=0.75$, for the Mini, SWA and 7D-PAR, respectively). Bland-Altman plots for estimates of total energy expenditure are shown in Figure 2. Correlations between the residuals and the DLW were

low for the Mini ($r=0.10$, panel a), the SWA ($r=0.04$, panel b) and the 7D-PAR ($r=0.01$, panel c). This indicates that there was no evidence of systematic bias in the estimations.

Estimates of EE from the Mini and the SWA were also reprocessed using the newly developed algorithms (algorithms v. 5.2). Again, group mean comparisons did not yield differences in TEE between the DLW method and both the Mini [diff= $6.9 \text{ kcal}\cdot\text{day}^{-1}$ (0.3%), 95% CI=-112.3 to 126.0] and the SWA [diff= $146.9 \text{ kcal}\cdot\text{day}^{-1}$ (5.5%), 95% CI=-37.0 to 256.8]. Absolute error values remained low for both monitors (Mini=7.3%, SWA= 8.0%). Pearson product-moment correlations for TEE between the DLW and the monitor's new algorithms remained high in both cases ($r=0.88$ and $r= 0.89$, for the Mini and SWA, respectively). Bland-Altman plots for TEE estimates comparisons with the monitor's new algorithms (APPENDIX B-Figure 1) showed similar confidence intervals and some systematic bias for both monitors (Mini $r= 0.29$, SWA $r=0.42$), with trends of EE underestimation observed at increased TEE values.

Analyses of Activity energy expenditure

Activity energy expenditure (AEE) is the most variable component of TEE and it can be estimated by subtracting the contribution from the thermic effect of food (TEF) and resting metabolic rates (RMR) from TEE. The TEF was estimated to be 10% of each participant's TEE based on standard convention. The measured RMR accounted for 55.6%, 56.6%, 58.1% and 53.7% of the TEE for DLW, Mini, SWA and 7D-PAR, respectively. The associated estimates of AEE were 34.4%, 33.4%, 31.9% and 36.3% of TEE, respectively (Figure 3). Group mean comparisons did not present differences in

AEE between the DLW method and estimates from the Mini monitor [-10.4 kcal·day⁻¹ (-1.1%), 95%CI=-118.7 to 139.5], the SWA monitor [-69.9 kcal·day⁻¹ (-7.6%), 95% CI=-33.5 to 173.3] and the 7D-PAR instrument [121.4 kcal·day⁻¹ (13.2%), 95% CI=-320.6 to 77.9]. On average, the two monitors underestimated TEE while the 7D-PAR overestimated TEE. Absolute error values with the DLW method were 21.9%, 17.1% and 33.5%, for the Mini, SWA and 7D-PAR, respectively. Pearson product-moment correlations between the DLW method and the comparison methods for AEE were moderate to high for both activity monitors (Mini $r=0.67$, SWA $r=0.76$) but low for the 7D-PAR ($r=0.19$). Bland-Altman plots (Figure 4) showed low systematic bias trends for the Mini ($r=0.25$, panel a) and for the SWA ($r=0.27$, panel b) monitors, with some EE underestimation by the monitors observed at increased AEE values. On the other hand, the Bland-Altman plot of the 7D-PAR residuals showed evidence of systematic bias ($r=0.57$, panel c), with larger EE underestimation for higher AEE values.

Discussion

The main purpose of the present study was to evaluate the validity of the SWA (software version 6.1, algorithms 2.2) and the Mini (software version 7.0, algorithms 2.2) monitors for measuring total daily EE and activity EE in older adults under free-living conditions, using direct comparisons with the DLW method. Both activity monitors did not show significant differences for TEE, and showed similarly low mean absolute error rates (SWA= 6.5%, Mini= 8.4%) compared to the DLW method. The use of mean absolute error rates allowed us to account for the substantial individual variability

observed in the estimates, and to make direct comparisons with other studies. It is particularly noteworthy that the values observed for the two monitors are almost identical to values recently reported by our group in a similar study of young and middle-aged adults (12).

To our knowledge, no previous study has assessed the validity of the Mini monitor in older adults. The mean absolute error rates, correlation coefficients and Bland-Altman plots were very similar for both multi-sensor monitors in the current study. Collectively, the results suggest that the Mini and the SWA provide comparable performance in older adults.

In this study, we were also able to test newly developed Armband algorithms (version 5.2). Comparisons with the DLW did not show significant differences for both Mini and SWA for TEE and similar absolute error rates and correlation coefficients as observed with the currently available algorithms. However, Bland-Altman plots showed some systematic bias for TEE. Therefore, no significant improvements were observed with the use of the newer algorithms in this sample of older adults. Previous studies in adults and children with the SWA (8, 24) have demonstrated improvements in accuracy with updated algorithms. It is not clear why the performance was somewhat worse with the newer versions. It may be that the samples used to develop the new algorithms did not include older adults.

In previous studies, researchers have reported a tendency of the SWA monitor to underestimate EE at high PA intensities (12, 25, 26). In the current study, Bland-Altman plots of TEE estimates did not show systematic bias for the SWA ($r=0.04$) or the Mini

($r=0.10$). One of the reasons for the absence of systematic bias in the sample might be the absence of vigorous PA that could accentuate the differences in EE at those high intensities. The descriptive results indicate that, on average, older adults accumulate only 3 minutes of vigorous activity a day.

Consistent with a previous study in our laboratory in young and middle-aged adults (12), the AEE estimates from the SWA and Mini showed larger mean absolute error rates compared to TEE estimates. Activity EE, as the most variable contributor to TEE, is complex in nature and difficult to measure accurately (16). Activity EE represents roughly a third of TEE (i.e., 34.4% in our sample); therefore, small errors in estimation of individual activities can have a marked impact on the overall accuracy of the assessment.

In a recent study in older adults (27), researchers compared five different PA assessment tools with the DLW method. In the study, an older version of the SWA (software version v.5.12) significantly underestimated PAEE ($p<0.008$) compared to the DLW method. The researchers reported a mean absolute underestimation of 26.8% and a moderate association with DLW ($r=0.48$) for PAEE. It is important to mention that in that previous study, the description of the methods suggests that the research group compared AEE from the DLW method with estimates of physical activity EE (PAEE) from the SWA, which reflects only the contribution of moderate and vigorous activity. The estimate of AEE from the DLW included light, moderate and vigorous activity so it is not surprising that this comparison would result in a significant underestimation of EE. In the current study, the differences with the DLW method were non-significant for the SWA

($p=0.17$) and for the Mini monitor ($p=0.87$). The mean absolute error for AEE was smaller for the SWA (22.0%) and the Mini (19.6%), and the correlations were high for both instruments. In addition, Bland-Altman plots for AEE in the current study showed similar trends of underestimation at higher intensities of EE, but less evidence of systematic bias. The present results show more favorable results than this previous study.

A secondary purpose of this study was to evaluate the validity of the 7D-PAR against the temporally matched data from the monitors and the DLW method. Self-report instruments provide a simpler and less expensive way to assess PA and they also provide information about the type, context and purpose of the activities performed. Self-report instruments are limited by challenges with recall, bias and social desirability (28). The comparison with the SWA data reveals clear discrepancies between the reported minutes of PA with the 7D-PAR and the activity monitors. The 7D-PAR underestimated time spent in moderate intensity PA and overestimated time spent in vigorous PA compared to the activity monitors. The underestimation in moderate intensity PA by the 7D-PAR is likely explained by the fact that the activity monitors record all the minutes of PA above a certain threshold while the 7D-PAR required that bouts be 10 minutes or longer. This would cause observed minutes of moderate PA to be higher than reported minutes. The overestimation of time spent in vigorous intensity PA by the 7D-PAR could be attributable to differences between the absolute intensity and the relative or perceived intensity by the participant (29).

The 7D-PAR instrument did not show groups differences for TEE, however, it showed a larger mean absolute error rate and a lower correlation with the DLW method,

compared to the activity monitors. Bland-Altman plots for TEE showed a similar systematic bias than the activity monitors, with larger confidence intervals. For AEE comparisons between the 7D-PAR and the DLW method, the 7D-PAR did not show significant group differences. On the other hand, the mean absolute error rate was very high (33.5%). Furthermore, the correlation coefficient between the self-report and the DLW method was very low and the Bland-Altman plot showed large systematic bias ($r=0.61$), with large underestimation at increased AEE. These results suggest larger individual variability, and limited accuracy compared to the activity monitors.

A previous study by Washburn and colleagues compared the 7D-PAR against the DLW method in young adults. In that study, the 7D-PAR provided non-significantly different estimates of mean TEE and PAEE (18). However, as observed in the current study, the 7D-PAR showed low agreement with the DLW ($r=0.37$) and individual estimates of EE showed considerable error. Previous studies in older adults comparing PA questionnaires to the DLW method have also shown similar low associations between the measures (30, 31, 32). In the study by Colbert and colleagues (27) using DLW in older adults, researchers compared PAEE estimates from a multisensor monitor (SWA v5.12), a commonly used accelerometry-based activity monitor (Actigraph GT1), a pedometer, and three self-report instruments (Yale Physical Activity Survey, Community Health Activities Model Program for Seniors (CHAMPS), and a modified Physical Activity Scale for the Elderly (modPASE). The results showed large PAEE mean absolute error rates for the objective measurements (range: 22.5-26.8%) but larger absolute error for the self-report instruments (range: 30.4-32.8%). Additionally, the

associations between the monitors and the DLW method were moderate (range: $r=0.48-0.60$), while the associations for the self-report instruments were low (range: $0.28-0.07$).

On average, the 7D-PAR overestimated TEE and AEE for our sample of older adults (non-significant differences). In a previous study, comparisons between self-report instruments and the DLW method (2), showed similar overestimations in PAEE and minutes of PA. Researchers suggested a possible influence of social desirability and social approval explaining the “over reporting” by the participants. An alternative explanations for the overestimation in EE with the 7D-PAR could be the use of standard metabolic equivalents ($MET= 3.5 \text{ mL O}_2 \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ or $1 \text{ kcal} \cdot \text{kg}^{-1} \cdot \text{hr}^{-1}$), which do not account for individual variability in RMR (33). A third possible explanation could be the characteristics of the 7D-PAR, that does not discriminate sitting time ($\sim 1-1.2$ METs) from light activity (1.5 METs), overestimating EE for every minute of sitting time (34).

The results from the study provide continued support for the use of the SWA and the Mini. A key goal in the project was to also attempt to better understand sources of error that may impact the accuracy in older adults. In the current study, the self-report instrument provided information about the types of activities that participants performed. If activity profiles are different in individuals with lower levels of agreement (compared to high) it might provide insights about factors that contribute error in estimation. In our study, we observed that participants that spent a significant amount of time driving a motor vehicle, engaged in cycling (stationary or road) and driving a motorcycle had above average error values. Previous studies have shown limitations of the SWA to assess EE estimates during cycling in adults (8) and children (24, 35) with

underestimations ranging from 12.0-28.9%. To our knowledge, no previous study assessed the validity of the SWA (or Mini) to estimate the energy cost of driving a car or a motorbike.

It is also important to consider other potential sources of error that may influence the estimations provided by the SWA and Mini monitors. In a recent study by Heiermann and colleagues (36), researchers reported reliable estimation of REE with the SWA (version 5.0) in older individuals (60-87 years old). However, results showed an overestimation of REE by the SWA monitor compared to indirect calorimetry (12-14% difference). Researchers suggested that the differences in estimation could be explained by age-related changes in skin conductance and thermoregulation, variables used by the monitor to estimate EE.

A previous study with the SWA showed significant differences in EE estimation when assessing EE in morbidly obese individuals (37), with underestimation in resting EE (8.8%) and overestimations in treadmill walking (31.4%). When assessing the effect of BMI on TEE estimation in our study, we observed that those individuals categorized as overweight (10.1%) or obese (11.1%) had higher average absolute error rates for the SWA monitor, compared with the normal BMI individuals (6.0%). In contrast, the Mini monitor did not show such large discrepancies when comparing BMI groups (5.7%, 8.4% and 7.7%, for normal, overweight and obese individuals, respectively), perhaps demonstrating an advantage of the Mini monitor over the SWA monitor when assessing EE in overweight and obese individuals. However, the small and unbalanced sample of participants in each BMI group for the current study precludes definitive conclusions.

In conclusion, the results indicate that SenseWear Pro₃ and SenseWear Mini Armband monitors provide valid estimates of TEE and AEE in older adults, under free-living conditions. The study also demonstrated reasonable group level estimates from the 7D-PAR but questionable accuracy for individual estimation. A strength of the study is the use of the DLW method, considered the “gold standard” for measuring EE under free-living conditions. The inclusion of measured RMR obtained via indirect calorimetry was important for reducing error in the AEE comparisons. A noteworthy aspect of the study is the high level of compliance noted for wear time (average of 97.9% wear time across days and participants). This may have contributed to the more favorable findings than reported by (27) in a similar study with older adults (average of 58.3% or 14 hours per monitoring day).

It is also important to acknowledge the limitations of the study. One of the limitations, is its small sample size due to the inherent limitations of the DLW method (i.e., costs, cumbersomeness, time commitment). The use of a convenience sample is another clear limitation of the study. The nature of the study makes it difficult to include a representative sample of the population of choice, thereby, precluding the generalization of the results to the general population. Additional work is needed to continue to advance research on physical activity assessment techniques in this population.

Acknowledgements

The authors gratefully acknowledge the enthusiastic support of the volunteers who participated in this study. No conflict of interest is reported by the authors.

Disclosure Statement

This research was funded by an NIH grant (HL091024-04) awarded to Dr. Greg Welk.

References

1- Sherman SE, D'Agostino RB, Cobb JL, Kannel WB. Does exercise reduce mortality rates in the elderly? Experience from the Framingham Heart Study. *Am Heart J.* 1994; 128: 965-972.

2- Adams SA, Matthews CE, Ebbeling CB, et al. The effect of social desirability and social approval on self-reports of physical activity. *Am J Epidemiol.* 2005; 161: 389–98.

3- Hendelman D, Miller K, Bagget C, Debold E, Freedson P. Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Med Sci Sports Exerc.* 2000; 32: S442–9.

4- Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System. United States, 2009.

5- Corder K, Brage S, Mattocks C, et al. Comparison of two methods to assess PAEE during six activities in children. *Med Sci Sports Exerc.* 2007; 39(12): 2180–8.

6- Welk GJ, McClain JJ, Eisenmann JC, Wickel EE. Field validation of the MTI Actigraph and BodyMedia armband monitor using the IDEEA monitor. *Obesity* (Silver Spring) 2007; 15(4): 918-928.

7- Fruin ML, Rankin JW. Validity of a multi-sensor armband in estimating rest and exercise energy expenditure. *Med Sci Sports Exerc.* 2004; 36(6): 1063-1069.

8- Jakicic JM, Marcus M, Gallagher KI, Randall C, Thomas E, Goss FL, Robertson RJ. Evaluation of the SenseWear Pro Armband to assess energy expenditure during exercise. *Med Sci Sports Exerc.* 2004; 36(5): 897-904.

9- King GA, Torres N, Potter C, Brooks TJ, Coleman KJ. Comparison of activity monitors to estimate energy cost of treadmill exercise. *Med Sci Sports Exerc.* 2004; 36(7): 1244-1251.

10- Malavolti M, Pietrobelli A, Dugoni M, Poli M, Romagnoli E, De Cristofaro P, Battistini NC. A new device for measuring resting energy expenditure (REE) in healthy subjects. *Nutr Metab Cardiovasc Dis.* 2007; 17(5): 338-343.

11- Cereda E, Turrini M, Ciapanna D, Marbello L, Pietrobelli A, Corradi E. Assessing energy expenditure in cancer patients: A pilot validation of a new wearable device. *J Parenter Enteral Nutr.* 2007; 31(6): 502-507.

12- Johannsen DL, Calabro MA, Stewart J, Franke W, Rood JC, Welk GJ. Accuracy of armband monitors for measuring daily energy expenditure in healthy adults. *Med Sci Sports Exerc.* 2010; 42(11): 2134-40.

- 13- Montoye HJ, Kemper HC, Saris WHM, and Washburn RA. *Measuring Physical Activity and Energy Expenditure*. Champaign, IL: Human Kinetics, 1996, pp. 17-25.
- 14- Starling RD. Use of doubly labeled water and indirect calorimetry to assess physical activity. In: Welk GJ. *Physical Activity Assessments for Health-Related Research*. Champaign, Ill: Human Kinetics; 2002: 197-209.
- 15- Westerterp KR, Brouns F, Saris WH, and ten Hoor F. Comparison of doubly labeled water with respirometry at low- and high-activity levels. *J Appl Physiol*. 1988; 65: 53-56.
- 16- Levine JA. Nonexercise activity thermogenesis (NEAT): environment and biology. *Am J Physiol Endocrinol Metab*. 2004; 286: E675-685.
- 17- Jones RH, Sonko BJ, Miller LV, Thureen PJ, Fennessey PJ. Estimation of doubly labeled water energy expenditure with confidence intervals. *Am J Physiol Endocrinol Metab*. 2000; 278: E383–E389.
- 18- Washburn, RA, Jacobsen DJ, Sonko BJ, Hill JO, and Donnelly JE. The validity of the Stanford seven-day physical activity recall in young adults. *Med. Sci Sports Exerc*. 2003; 35(8): 1374-1380.

19- Sallis J F, Haskell W, Wood P, et al. Physical activity assessment methodology in the Five-City Project. *Am J Epidemiol.* 1985; 121:91-106.

20- Schoeller DA. Measurement of energy expenditure in free-living humans by using doubly labeled water. *J Nutr.* 1988; 118: 1278-1289.

21- Racette SB, Schoeller DA, Luke AH, Shay K, Hnilicka J, and Kushner RF. Relative dilution spaces of 2H- and 18O-labeled water in humans. *Am J Physiol.* 1994; 267: E585-590.

22- Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, O'Brien WL, Bassett DR Jr, Schmitz KH, Emplaincourt PO, Jacobs DR Jr, Leon AS. Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc.* 2000; 32: S498-504.

23- Bland JM and Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986; 1: 307-310.

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

25- Drenowatz C, Eisenmann JC. Validation of the SenseWear Armband at high intensity exercise. *Eur J Appl Physiol*. 2010. [Epub ahead of print]

26- Koehler K, Braun H, de Marées M, Fusch G, Fusch C, Schaenzer W. Assessing energy expenditure in male endurance athletes: validity of the Sensewear Armband. *Med Sci Sports Exerc*. 2010. [Epub ahead of print]

27- Colbert LH, Matthews CE, Havighurst TC, Kim K, and Schoeller DA. Comparative validity of physical activity measures in older adults. *Med Sci Sports Exerc*. 2010. Publish ahead of print.

28- Matthews CE. Use of self-report instruments to assess physical activity. In: Welk GJ, ed. *Physical Activity Assessments for Health-Related Research*. Champaign, IL: Human Kinetics; 2002:107–123.

29- Altschuler A, Picchi T, Nelson M, Rogers JD, Hart J, and Sternfeld B. Physical activity questionnaire comprehension: lessons from cognitive interviews. *Med Sci Sports Exerc*. 2009; 41(2): 336–343.

30- Bonnefoy M, Normand S, Pachiardi C, Lacour J, Laville M, Kostka T. Simultaneous validation of ten physical activity questionnaires in older men: a doubly labeled water study. *J Am Geriatr Soc*. 2001; 49:28-35.

- 31- Mahabir S, Baer D, Giffen C, Clevidence B, Campbell W, Taylor P, Hartman T. Comparison of energy expenditure estimates from 4 physical activity questionnaires with doubly labeled water estimates in postmenopausal women. *Am J Clin Nutr.* 2006; 84:230-6.
- 32- Starling RD, Matthews D, Ades P, Poehlman E. Assessment of physical activity in older individuals: a doubly labeled water study. *J Appl Physiol.* 1999; 86: 2090-6.
- 33- Byrne NM, Hills AP, Hunter GR, Weinsier RL, Schutz Y. Metabolic equivalent: one size does not fit all. *J Appl Physiol.* 2005; 99: 1112–1119.
- 34- Welk GJ, Thompson RW, Gaper DI. Temporal validation of scoring algorithms for the 7-day physical activity recall. *Measurement in Physical Education & Exercise Science.* 2001; 5(3) 3, 123-138.
- 35- Arvidsson D, Slinde F, Larsson S, Hulthen L. Energy cost of physical activities in children: validation of SenseWear Armband. *Med Sci Sports Exerc.* 2007; 39(11): 2076-2084.
- 36- Heiermann S., Hedayati KK, Müller MJ, Dittmar M. Accuracy of a portable multisensor body monitor for predicting resting energy expenditure in older people: a comparison with indirect calorimetry. *Gerontology.* 2010. [Epub ahead of print]

37- Papazoglou D, Augello G, Tagliaferri M, Savia G, Marzullo P, Maltezos E, Liuzzi A. Evaluation of a multisensor armband in estimating energy expenditure in obese individuals. *Obesity* (Silver Spring) 2006; 14(12): 2217-2223.

Tables

Table 1. Representation of the experimental protocol.

Day	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Body weight recorded	▲							▲							▲
$^2\text{H}_2^{18}\text{O}$ dosage	▲														
Urine collection	▲ x6							▲ x2							▲ x2
Mini on body	→														
SWA on body	→														
Resting metabolic rate								▲							▲
7-day PA recall								▲							▲

Table 2. Descriptive statistics for the sample population.

	All		Men		Women	
	Mean	SD	Mean	SD	Mean	SD
Age (years)	67.9	7.0	67.6	6.6	68.1	7.6
Height (cm)	169.3	7.4	173.9	7.0	165.4	5.3
Weight (kg)	77.2	19.3	82.6	22.5	72.8	16.0
BMI (kg/m²)	26.7	5.3	27.1	5.8	26.3	5.1
RMR (kcal·day⁻¹)	1485.0	409.4	1683.9	457.3	1322.3	294.0
TEF (kcal·day⁻¹)	267.0	52.0	306.2	46.9	235.0	29.3
PAL	1.8	0.3	1.9	0.2	1.8	0.3

Table 3. Total daily energy expenditure values (kcal·day⁻¹) for each assessment tool.

	All		Men		Women	
	Mean	SD	Mean	SD	Mean	SD
DLW	2635.6	509.4	2691.9	489.7	2349.9	293.3
Mini	2624.0	561.9	2701.9	571.6	2321.7	317.5
SWA	2557.9	572.7	2661.0	577.0	2241.0	308.0
7D-PAR	2770.4	688.9	2713.5	812.0	2565.1	531.6

Abbreviations: Doubly labeled water (DLW), SenseWear Mini Armband (Mini), SenseWear Pro₃ Armband (SWA) and 7-day physical activity recall (7D-PAR).

Figures

Figure 1- Moderate and vigorous physical activity contribution by assessment tool (minutes).

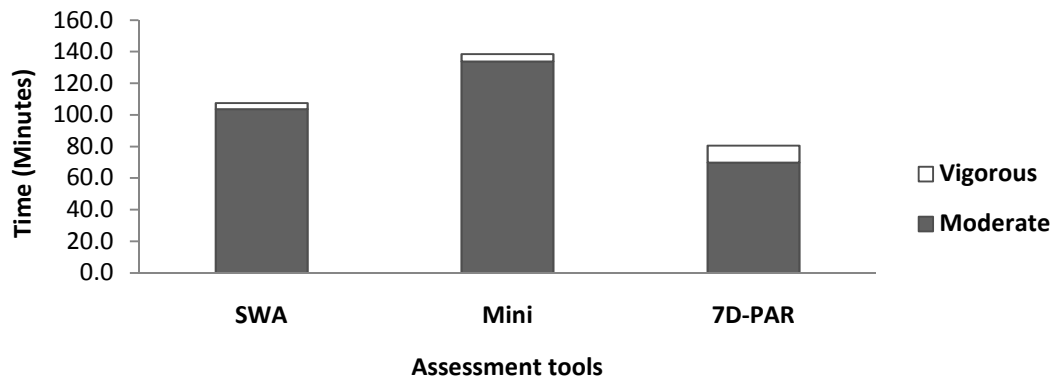
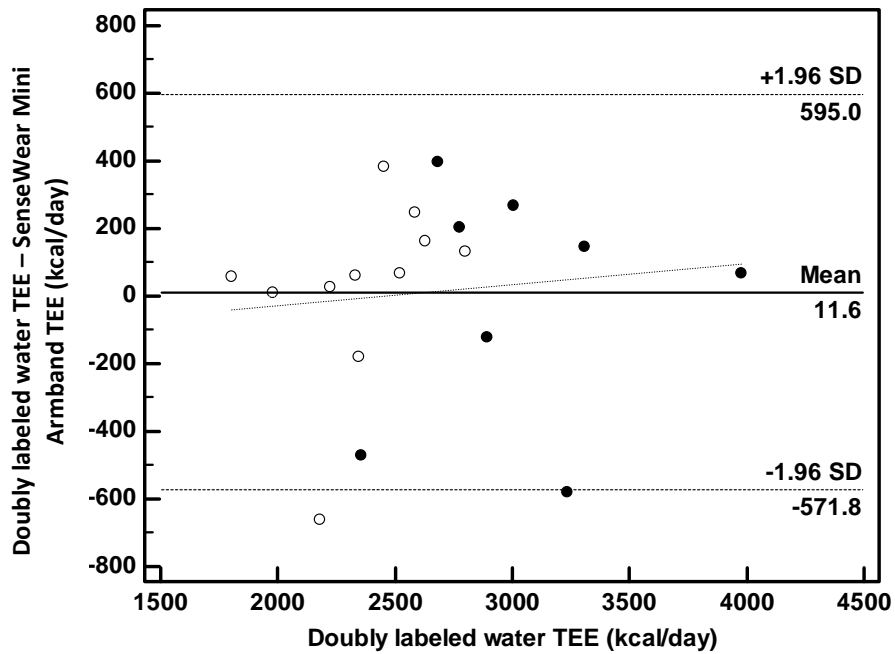


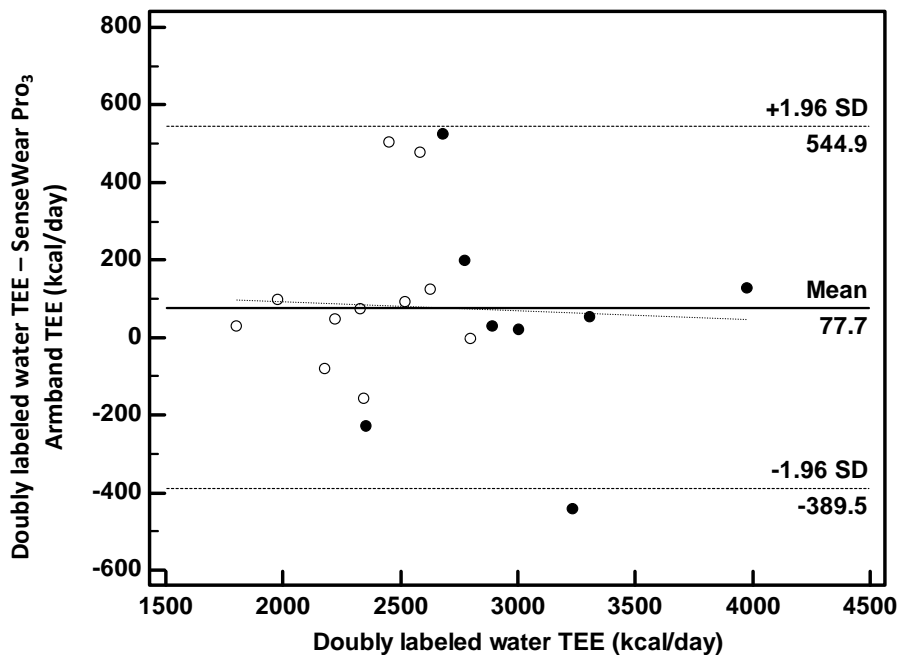
Figure 2- Bland Altman plots for total energy expenditure ($\text{kcal}\cdot\text{day}^{-1}$)

Gender reference: ● Males, ○ Females

a) SenseWear Mini Armband (Mini)



b) SenseWear Pro3 Armband (SWA)



c) 7 Day physical activity recall (7D-PAR)

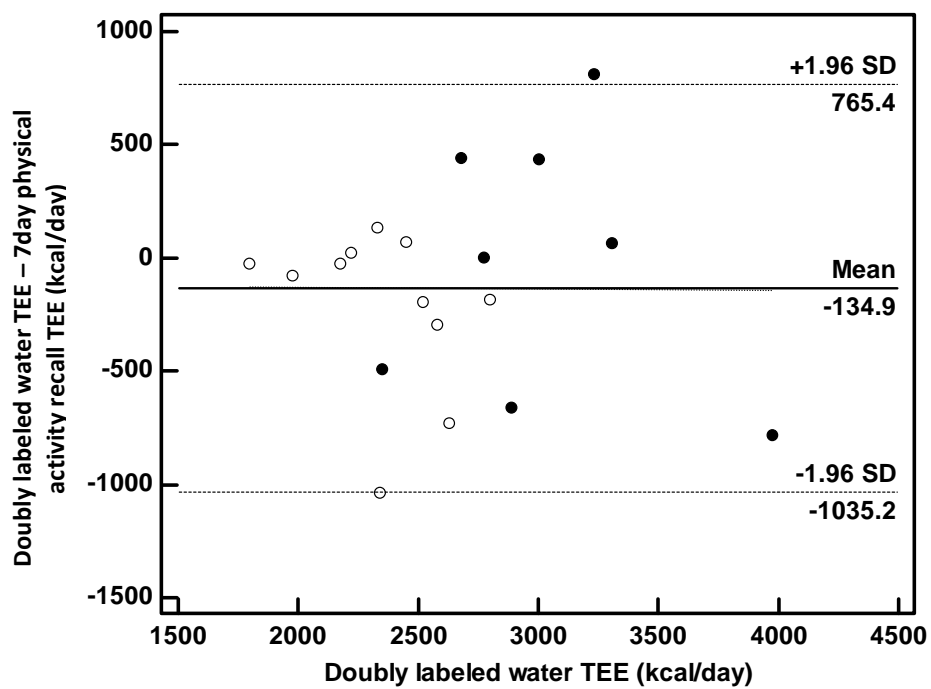
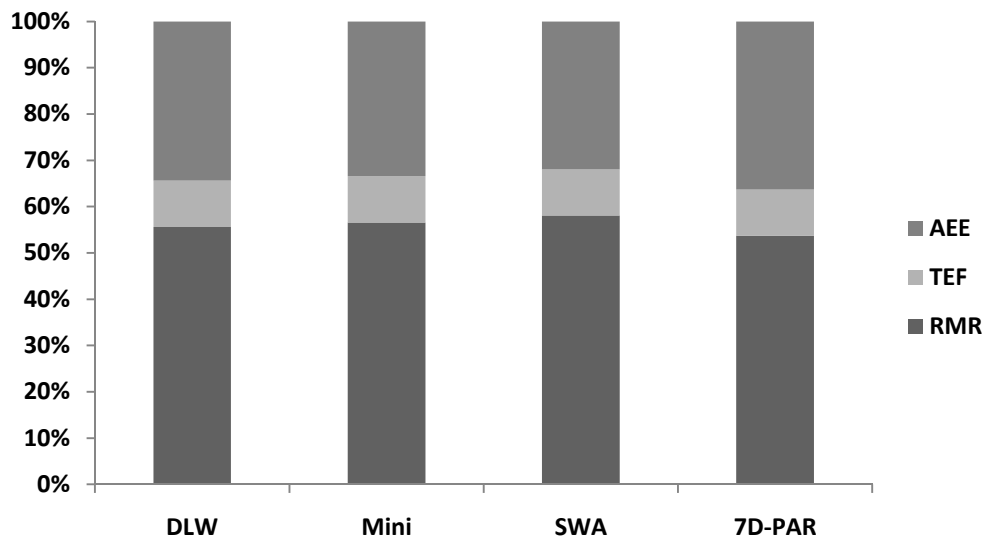


Figure 3- Total energy expenditure distribution by assessment tool.

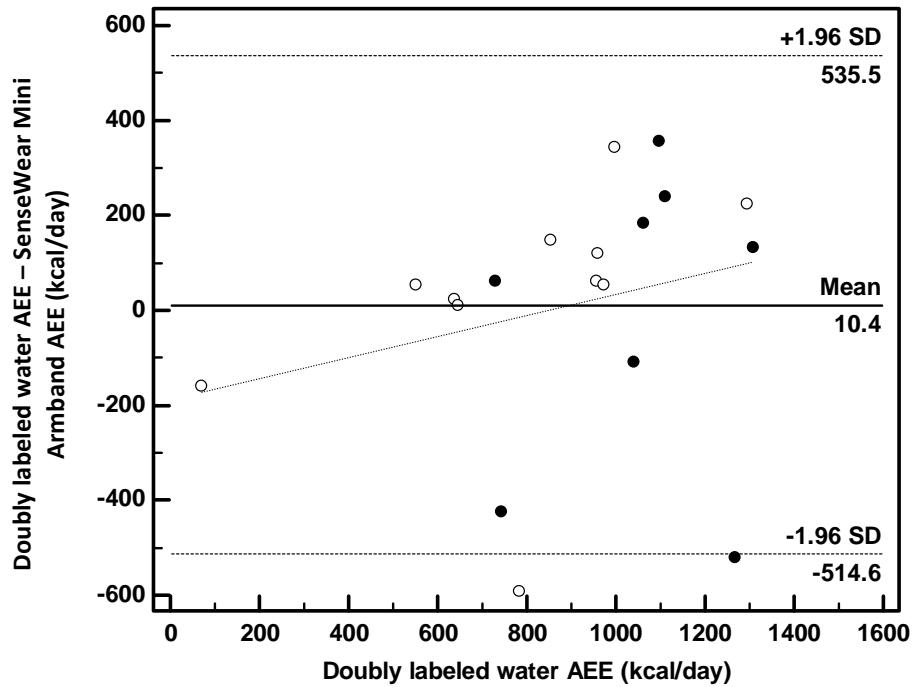


Abbreviations: Doubly labeled water (DLW), SenseWear Mini Armband (Mini), SenseWear Pro₃ Armband (SWA) and 7-day physical activity recall (7D-PAR). Activity energy expenditure (AEE), thermic effect of foods (TEF) and resting metabolic rate (RMR).

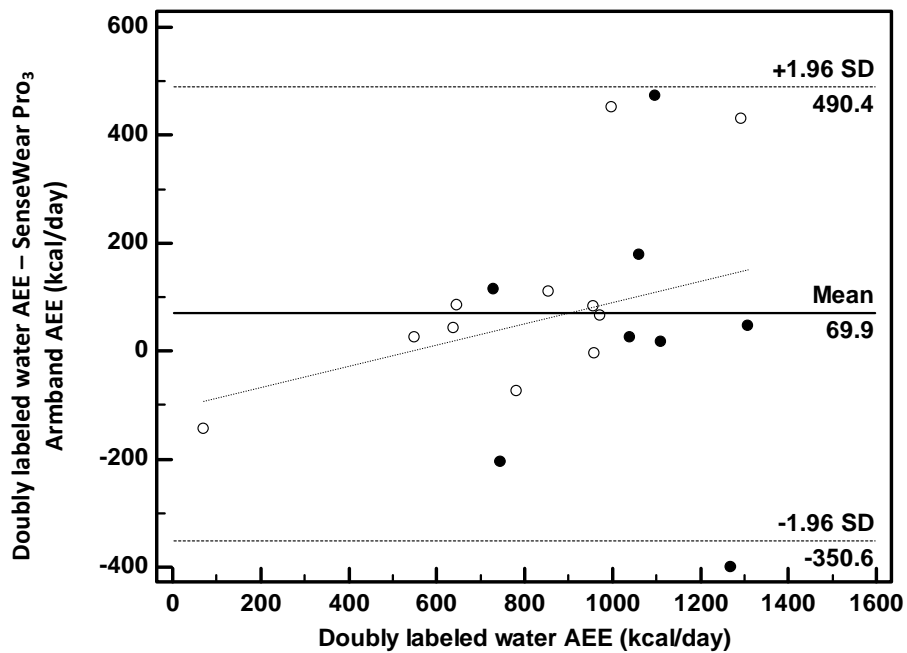
Figure 4- Bland Altman plots for activity energy expenditure (AEE, kcal.day⁻¹)

Gender reference: ● Males, ○ Females

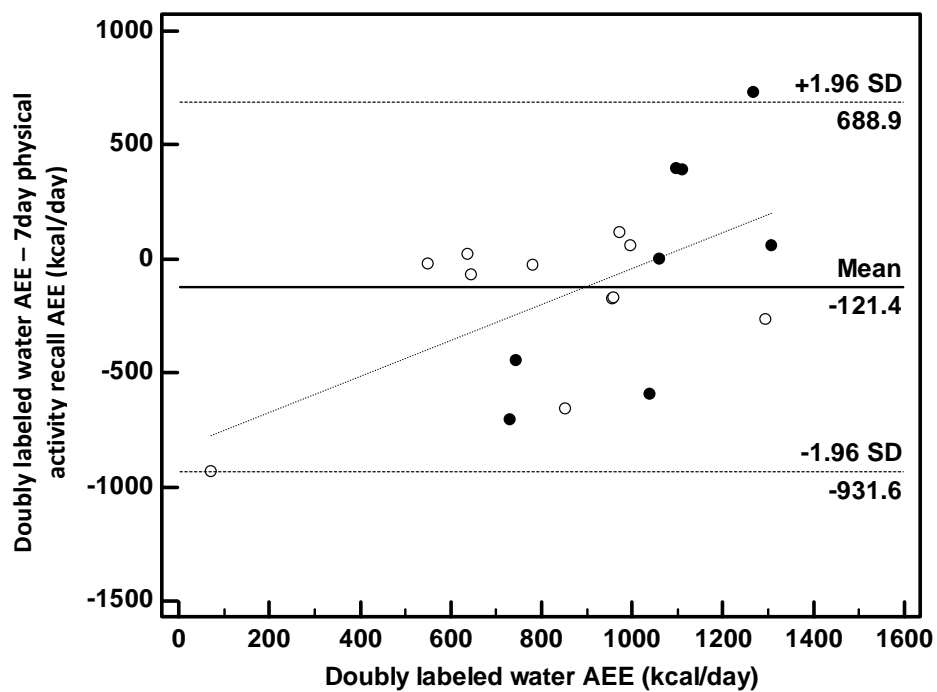
a) SenseWear Armband Mini (Mini)



b) SenseWear Pro3 Armband



c) 7-Day physical activity recall (7D-PAR)



CHAPTER 4. VALIDATION OF PATTERN-RECOGNITION MONITORS IN CHILDREN USING DOUBLY LABELED WATER.

A manuscript to be submitted for publication in *Medicine and Science in Sports and Exercise*

M. A. Calabró, J.M. Stewart, G. J. Welk.

Abstract

Accurate assessments of physical activity and energy expenditure are needed to advance research on childhood obesity prevention. **Purpose:** The purpose of the study was to evaluate the validity of two SenseWear Armband monitors [The SenseWear Pro3 (SWA)] and the recently released Sensewear Mini (Mini) under free-living conditions in a youth population. **Methods:** Participants in the study (28 healthy children aged 10-16 yr) wore both monitors for 14 consecutive days, including sleeping time. Estimates of total energy expenditure (TEE) from the SenseWear monitors were computed using two different algorithms (version 2.2 algorithms, available in the SenseWear software version 6.1 and 7.0, and the newly developed 5.0 algorithms). The estimates were compared to estimates derived from the doubly labeled water methodology (DLW) using standard measurement agreement procedures. **Results:** The refined 5.0 algorithms did not yield significant differences for the SWA [123.3 (3.8%) kcal·day⁻¹] but showed significant differences for the Mini [206.0 (6.9%) kcal·day⁻¹, p<0.001]. The absolute error rates (computed as average absolute value of the individual errors) for the 2.2 algorithms were more similar (10.4% for the SWA 11.4% for the Mini). Pearson product-moment correlations were

high for all monitors/algorithms [correlation coefficients (r) were 0.93, 0.92, 0.90 and 0.90, for the SWA2.2, Mini2.2, SWA5.0 and Mini5.0, respectively]. **Conclusions:** The newly developed SenseWear Armband 5.0 algorithms outperformed the version 2.2 algorithms. The low absolute error rates of the SWA and Mini monitor support the validity of the new algorithms for assessing TEE under free-living conditions in youth.

Key words: Energy expenditure, physical activity, activity monitor.

Introduction

There is considerable public health interest in assessing and promoting physical activity (PA) in children. Accurate measures are needed to evaluate population patterns and trends, for understanding correlates of activity behavior, and for evaluating health outcomes and interventions. An accurate measurement of daily energy expenditure (EE) under free-living conditions is especially important for understanding how PA contributes to overweight and obesity. Because self-report instruments have acknowledged limitations in youth (Slootmaker et al. 2009), emphasis has been placed on objective assessment techniques such as the use of accelerometry-based activity monitors. Considerable research has been conducted with a variety of monitors but research has demonstrated clear limitations associated with the use of standard, uni-axial accelerometers (Trost et al. 2005). Standard accelerometry-based devices work reasonably well for locomotor-based activities but are not well suited to capturing the diverse range of lower intensity lifestyle activities that comprise the bulk of the day. Challenges associated with assessing compliance and in calibrating monitors for different ages have proven particularly difficult to resolve. The assessment of PA and EE in children is further confounded by variable activity patterns and variability in energy cost of activity due to growth and maturation (Harrell et al. 2005).

Multi-sensor pattern-recognition monitors offer considerable promise for improving estimates of PA and EE. The combination of multiple sensors allows for the detection of PA patterns and the application of movement specific algorithms to estimate

EE. Pattern-recognition monitors have been shown to provide more accurate estimates of PA than commonly used accelerometers (Corder et al. 2007; Welk, 2007).

The SenseWear Pro3 Armband (SWA, BodyMedia Inc., Pittsburgh, PA) is a wireless, multi-sensor monitor that integrates information from a 2-axis accelerometer with a variety of heat-related sensors (i.e., heat flux, skin temperature, near-body ambient temperature, and galvanic skin response). The integration of heat-related sensors allows the SWA to estimate the energy cost of complex movements and upper body activities that have proven to be difficult to assess with hip-worn accelerometers.

Previous studies have evaluated the validity of the SWA for estimation of EE in children under laboratory conditions (Arvidsson et al. 2007; Arvidsson et al. 2009a; Dorminy et al. 2008; Calabro et al. 2009). However, to date, only one study has evaluated the validity of the SWA in children under free-living conditions (Arvidsson et al. 2009b). In that study, researchers compared TEE estimates from two versions of the software of the SWA (Software v.5.1 and v.6.1) with the doubly labeled water method (DLW, the gold standard for TEE assessment under free-living conditions). Researchers reported that the current software (v.6.1, algorithm version 2.2) improved the accuracy for group level estimation, but individual error was considerable and dependent on physical activity level. Appropriately, a newer set of children algorithms (5.0), to be proximally released by the manufacturer, have not been yet evaluated under free-living conditions. The recently developed SWA Mini (Mini, Software v.7.0, algorithms version 2.2) has also not been evaluated for use in children. A recent DLW study in adults (Johannsen et al. 2010) demonstrated that the Mini yielded more accurate estimates of EE than the SWA in adults

(possibly because of the use of a 3-dimensional accelerometer in the unit). The present study evaluates the validity of the SWA and the Mini for measuring total daily EE in children using a similar approach. The study also directly compares the relative accuracy of the previous algorithms (version 2.2) with the newly developed children algorithms (version 5.0) to determine if the algorithms have also improved.

Methods

A sample of 30 healthy youth (age range: 10-16 yr) were recruited to participate in the study. A targeted recruitment process was used to ensure that participants were able to comply with the measurement protocol. While this reduced the generalizability of the sample population, it was determined to be more important to ensure compliant participants (i.e., internal validity was prioritized over external validity). Approval from the Institutional Review Board was obtained before the beginning of the study to ensure that the procedures were appropriate for the target population. All participants and their parents were informed about the procedures and purposes of the study before parental consent and participant assent were obtained.

Instruments

The *SenseWear Pro3 Armband (SWA, Model 908901 PROD2)* is a wireless multi-sensor activity monitor that integrates motion data from two orthogonal accelerometers along with several heat related sensors (heat flux, body temperature and galvanic skin response). The monitor, worn on the upper arm (right side) over the triceps muscle, is lightweight (79-grams) and comfortable to wear. The SWA has been

previously validated in adults (St-Onge et al. 2007; Johannsen et al. 2010) and children (Arvidsson et al. 2009b) under free-living conditions. The data obtained from the SWA was processed using the latest proprietary algorithms available in the software (Research Software 6.1, algorithms 2.2).

The *SenseWear Mini (Mini, Model MS-SW)* is a newer and smaller version of the SWA that is worn on the left arm. The Mini operates in a similar manner as the more established SWA but includes a tri-axial accelerometer instead of a two axis accelerometer. Data obtained from the Mini was processed using the same proprietary algorithms, but from a different software package (Research Software 7.0, algorithms 2.2).

Data Collection Procedures

Table 1 provides a schematic representation of the experimental protocol. On the first day of the study (day 0), participants reported to the campus research center following a 10-hour overnight fast (no food or drink other than water) and after collecting a baseline urine sample (Baseline A). Participants then provided a second baseline urine sample upon arrival (Baseline B) and standard anthropometric data were collected. Standing and sitting height were measured to the nearest 0.1 cm with the use of a wall mounted Ayrton stadiometer (Prior Lake, MN) and with the participants barefoot. Body mass was measured with participants in light clothes and barefoot on a Cardinal Detecto electronic scale (Webb City, MO) to the nearest 0.1 kg. The body mass index (BMI) was calculated as $\text{weight (kg)/height}^2 \text{ (m}^2\text{)}$. In addition, sitting height was measured following

standard procedures (Ross et al. 1991) in order to predict age at peak height velocity using equations developed by Mirwald and colleagues (2002).

The two monitors were initialized using the participant's personal information (age, gender, height, weight) and adjusted to fit on the participants arms. The SWA monitor was placed on the right arm while the Mini monitor was placed on the left arm, following manufacturer recommendations. After fitting both activity monitors, the DLW dose was administered to the participant. The dosage was determined based on body weight (1.5g per kg of body weight) in accordance with the protocol. Regular measured water was provided to clean the drinking container and ensure that all the "heavy water" was consumed by the participant.

The DLW procedure involved collection and processing of urine samples on days 0, 7 and 14. On Day 0, participants provided urine samples at 1.5, 3.0, 4.5 and 6 hours and returned samples to the research center later in the day. Participants were provided with a cooler bag containing pre-labeled 60 ml sterile cups and were asked to provide at least 40 ml in each sample. Participants were given a liter of fresh drinking water immediately upon administration of the dose and were encouraged to drink the water over the course of the morning to ensure adequate urine volume.

On Day 7 and Day 14, participants reported to the laboratory after a 10-hr fast and were asked to provide additional urine samples at two time points (90 minutes apart). Body weight was measured on both days to check for changes in weight over the course of the study. Resting metabolic rate (RMR) data was obtained on Day 7 between the collection of the two urine samples. These measurements were obtained in an isolated

dark room, with participants awake, in a reclined position, avoiding speaking and minimizing their movement, while watching a movie, using a metabolic measurement system (True One 2400[®], Parvo-Medics Inc., Sandy, UT). Pediatric size masks (Hans Rudolph Inc, Kansas City, Missouri) were fitted to the participants and properly adjusted before data collection. The RMR values were computed on a per minute basis and expressed per day to facilitate analyses. The first 10 minutes of data collected were discarded and the last 15 minutes averaged to obtain an estimate of 24-h RMR. The temperature of the room was maintained at 22° C and the metabolic analyzer was calibrated before every measurement for pressure and gas concentrations.

All urine samples were processed by the same research technician using standardized procedures. The specimens were labeled and coded by time to ensure accurate processing of the data. Duplicate urine samples of approximately 12ml were stored frozen and later sent for processing at the Pennington Biomedical Research Center (Baton Rouge, LA).

Total energy expenditure (TEE) was determined from DLW over a 14-day period by tracking the relative loss of the labeled isotopes (²H deuterium and ¹⁸O) in the water. The difference between the rates of disappearance of the isotopes reflects the total carbon dioxide (CO₂) production over the measured period and is calculated from the slope of the elimination curve. The rates of disappearance were determined from the multiple urine samples obtained throughout the protocol. A fixed respiratory quotient of 0.86 was used to establish oxygen consumption and to obtain a value for TEE over the 14 days.

Values were expressed per day to facilitate interpretation. Additional information about the DLW protocol and processing is published elsewhere (Johannsen et al. 2010).

Processing of SWA and Mini data

Participants in the study wore both devices simultaneously, following manufacturer recommendations, for the entire 14-day period (with the exception of showering time). During visits to the laboratory on Days 7 and 14 of the protocol, both monitors were downloaded for memory clearance and their batteries were recharged. All SWA and Mini files were processed with the latest version of the algorithms available for the latest software package (algorithms 2.2). In addition, all SWA and Mini raw files were sent to the manufacturer (BodyMedia, Inc.) to be processed with newly developed children algorithms (v. 5.0).

During their visits on Days 7 and 14, participants reported non-wearing periods (i.e., showering or dressing). Individual attention was given to each data file in order to control for possible gaps in the data during the monitoring period. Active non-wearing periods were compared with the reported non-wearing periods in order to account for possible gaps in the data. Those gaps were manually filled with corresponding MET values based on the Compendium of energy expenditures for youth Ridley et al. (2008). For example, data gaps attributable to showering were manually filled with a corresponding MET equivalent for “showering and toweling off” (2.0 METs) based on the Compendium. Other unaccounted gaps shorter than 10 minutes, commonly occurring due to a loose monitor strap, were filled with average EE of the 10 minutes before the gap. In addition, a group of 4 participants were involved in a volleyball league that

prevented them from using the monitors during tournament games. Those individuals recorded their playing time during volleyball practices with the activity monitors, and mean values of those periods were used to fill gaps produced during the tournament games.

Statistical analyses

Descriptive statistics (sample mean and standard deviations) were computed for the primary PA categories to describe the characteristics of the participants and their activity profiles. Data were checked for normality to ensure that the distribution of the data would not influence the results. Activity energy expenditure (AEE) and physical activity level (PAL) were calculated using the following equations: $AEE = (TEE \cdot 0.9) - RMR$ (assuming thermic effect of food to be 10% of TEE) and $PAL = TEE \cdot RMR^{-1}$, respectively.

The study evaluated the agreement between estimates of TEE and PAEE from the SWA and the Mini compared with criterion estimates from the DLW. Primary statistical analyses were performed using IBM SPSS Statistics 19 (IBM Corporation, Somers, NY). Statistical power was estimated for associated *t*-test at the 5% level. Paired *t* tests were used to determine differences between the mean values obtained with the SWA monitors (SWA and SWA Mini) and DLW. Analyses of covariance (ANCOVA) were used to test for the effects of gender and maturation on estimates of TEE. Furthermore, to evaluate the extent of agreement between measures of TEE and AEE, Pearson Product-Moment correlations were computed.

Bland Altman graphical procedures (Bland and Altman 1986) were used to examine agreement across the range of TEE and AEE values and evaluate the presence of systematic bias. The DLW values were plotted in the x-axis and the residuals between the estimates (i.e., DLW minus SWA) were plotted in the y-axis. Confidence intervals defining the limits of agreement were established as 1.96 SD from the mean difference. Differences were considered significant at $p < 0.05$.

Results

The main objective of the study was to compare measures of TEE between two activity monitors (Mini and SWA) and the DLW method, in a diverse sample of youth (age range 10-16). Thirty participants (15 male, 15 female) completed the study. The majority of the participants were Caucasian (76.0%) with 14.0 % Hispanic and 10.0 % Asian. There was a range of body types with approximately 16.7 % characterized as “at risk for overweight” (between 85th and 95th percentile), 3.3 % characterized as “overweight” (> 95th percentile), and 6.7 % characterized as “underweight” (< than 5th percentile). From the initial sample, two participant’s data had to be discarded from the analyses due to unusable DLW values (urine collection problems). In addition, the SWA monitor showed data abnormalities at downloading time in three trials and those SWA trials were not included in the analyses. Therefore, final analyses include data from 28 participants for DLW and the Mini monitor, and data from 25 participants using the SWA monitor.

On average, participants wore the monitors for 96.7% (± 3.0) of the time during the 14 days of monitoring. The participant's wearing percentages ranged from 85.6% to 99.9 % of protocol time. Furthermore, off-body time during the 14 days of monitoring was on average 336.4 minutes per week (± 320.5), and ranged from 9 to 1459 minutes. Analyses of covariance (ANCOVA) used to test for the effects of gender and maturation (years from age at peak height velocity) on estimates of TEE showed no significant effects for any of the comparisons with DLW; therefore, boys and girls were combined for all analyses. Additionally, no significant effect of maturation was found for any of the comparisons with DLW. Measured RMR for the sample was 1460.9 (± 367.3) kcal·day⁻¹ while the PAL was 1.81 (± 0.33). Descriptive statistics for the sample population are provided in Table 2. The statistical power to detect EE differences with the 28 participants in this sample was ~60%.

Total energy expenditure comparisons

Table 3 includes mean values for TEE (daily values) for DLW, as well as Armband monitor data for the currently available monitor algorithms (A) and the newly developed algorithms (B). Differences in TEE were significant ($p < 0.001$) when comparing the DLW method with the version 2.2 algorithms [-492.2 (20.4%) kcal·day⁻¹ and -365.6 (15.8%) kcal·day⁻¹ for the SWA and Mini, respectively]. Absolute error values were 20.7% and 18.3% for the SWA and Mini respectively. Comparisons between the DLW method and the newly developed algorithms (version 5.0) showed smaller differences in TEE estimation for both monitors. The difference for the SWA [123.3

(3.8%) kcal·day⁻¹] was non-significant (p=0.21) while the values for the Mini [206.0 (11.4%) kcal·day⁻¹] were significantly different (p=0.02). Absolute error values were similar for the SWA and Mini (10.4% and 11.4%, respectively). Pearson product-moment correlations between the DLW method TEE and the monitors were consistently high for both monitors, regardless of which algorithms were being used [correlation coefficients (r) with DLW were 0.93, 0.92, 0.90 and 0.90, for the SWA2.2, Mini2.2, SWA5.0 and Mini5.0, respectively].

Bland-Altman plots for TEE were used to assess systematic bias between the DLW and the monitor estimates (Figure 1). The mean TEE from the DLW method was plotted in the x-axis, while the differences between the monitors/algorithms and the DLW method were plotted in the y-axis. Limits of agreement from the plots (mean \pm 1.96 SD) were smaller for the version 2.2 algorithms (SWA: -34.3 to 1018.8 kcal·day⁻¹; Mini: -251.3 to 992.5 kcal·day⁻¹) compared to the version 5.0 algorithms (SWA: -887.2 to 640.6 kcal·day⁻¹; Mini: -971.6 to 559.7 kcal·day⁻¹). The plots showed a consistent underestimation of TEE for the version 2.2 algorithms, with the SWA underestimating TEE for 24 of the 25 participants (96.0%) and the Mini underestimating TEE for 25 of the 28 participants (89.3%). The plots for the version 5.0 algorithms were more balanced but showed a tendency to overestimate TEE, with the SWA overestimating TEE for 15 out of 25 participants (60%), and the Mini overestimating TEE for 16 of the 28 participants (57.1%). All TEE plots showed some form of systematic bias with overestimation larger for individuals with greater TEE values. The coefficients of

determination (R^2) for the systematic bias were: 0.05, 0.22, 0.19 and 0.22, for the SWA2.2, Mini2.2, SWA5.0 and Mini5.0, respectively (See Figure 1).

Activity energy expenditure

The mean AEE value for the DLW method was $859.4 \text{ kcal}\cdot\text{day}^{-1}$. Differences in AEE were significant ($p < 0.001$) when comparing the DLW method with the version 2.2 algorithms (-443.0 (51.5%) $\text{kcal}\cdot\text{day}^{-1}$ and -329.0 (38.2%) $\text{kcal}\cdot\text{day}^{-1}$ for the SWA and Mini, respectively). Absolute error values were 61.9% and 53.3% for the SWA and Mini, respectively. The differences in AEE estimates with the version 5.0 algorithms were considerably smaller (for both the SWA (111.0 (13.0%) $\text{kcal}\cdot\text{day}^{-1}$) and the Mini (185.4 (21.5%) $\text{kcal}\cdot\text{day}^{-1}$). The difference with the SWA were non-significant ($p = 0.13$) but the values for the Mini were significantly different ($p < 0.05$). Absolute error values for the SWA and Mini were lower with the 5.0 algorithms (31.5% and 32.5%, respectively). Pearson product-moment correlations between the DLW-derived AEE values and the monitor estimates were lower than for TEE, but still high for both monitors, and with both algorithms (correlation coefficients (r): 0.79, 0.83, 0.73 and 0.75, for the SWA2.2, Mini2.2, SWA5.0 and Mini5.0, respectively). Bland-Altman plots for AEE estimates were used to assess systematic bias (Figure 2) and showed similar trends as the TEE plots, with a similar form of systematic bias (tendency for greater overestimation for individuals with higher activity levels).

Discussion

The primary aim of the study was to assess the validity of the SWA (software version 6.1) and the Mini (software version 7.0) using DLW as the reference method, in a sample of youth under free-living conditions. The two monitors significantly underestimated TEE (20.4% and 15.8% for the SWA and Mini, respectively) compared to the DLW method. A secondary aim of the study was to test newly developed SWA and Mini children algorithms (v. 5.0) and compare them with the previous available version (v. 2.2). The newly developed 5.0 algorithms yielded more accurate estimates of TEE for both Armband monitors (3.8% and 11.4% for the SWA and Mini, respectively). There is considerable individual variability in the magnitude and direction of error so the absolute error rates provide a more appropriate indicator of true error. The values observed for the SWA and Mini (10-11%) are similar to values recently reported by our group in a similar study in adults (Johannsen et al. 2010).

Previous studies using the SWA in children have reported discrepant findings related to TEE estimation but the variability is likely due to differences in methods, samples and the versions/algorithms used in the study. Arvidsson and colleagues (et al. 2007) assessed the validity of EE estimates from the SWA in 20 healthy children (ages 11-13 yrs) using indirect calorimetry as the reference method, during a wide range of activities under laboratory conditions. They reported a significant underestimation in AEE by the SWA (v.5.1) for most activities (~22% on average). A later study by the same laboratory (Arvidsson et al. 2009a), using a comparable sample of children (mean age: 12.3 yrs) and the same method (indirect calorimetry), reported underestimations in

AEE of 18% for the SWA. The researchers in this study noted that the underestimation in AEE increased with the increasing intensity in the activities. Contrasting findings were reported for the SWA (v. 5.1) by Dorminy (et al. 2008) in a sample of youth (ages 10-14 yrs) monitored with indirect room calorimetry over a 24 h period. Researchers reported a consistent overestimation of EE for a variety of tasks (treadmill exercise, stationary biking, treadmill walking, sedentary activities) measured by whole-room indirect calorimetry. The overall error (overestimation) in TEE for the 24 hr. of monitoring was 22%. We reported a similar tendency for overestimation of EE (32%) with version 5.1 of the SWA software in a laboratory study in young children (ages 7-11 yrs) (Calabro et al. 2009). However, non-significant differences in EE were observed when we used the more recent version (6.1) of the software (average group level error of 1.7%).

The present study conducted with DLW as the criterion measure was conducted to help address these discrepancies in findings. The recently developed algorithms (5.0) clearly outperformed the version 2.2 algorithms that were used in the previous studies mentioned above. This proved to be true for both the SWA and the Mini. Interestingly, we observed slightly better performance for the SWA compared to the newer Mini despite the fact that the Mini uses a 3 dimensional accelerometer. A previous study in adults (Johannsen et al. 2010) demonstrated some improved accuracy for the Mini relative to the SWA, so additional work may be needed to understand these differences.

The results of the present study are similar to findings of another DLW study (Arvidsson et al., 2009b). This study demonstrated clear improvements with the version 6.1 algorithms compared to the previously available software version (5.1). Researchers

reported a significant mean overestimation (8.3%, $p < 0.01$) with the older version of the software and an improved non-significant estimation difference (6.0%) with the newer software version (v.6.1). Consistent with our findings, the researchers reported high Pearson product-moment correlations between the SWA and the DLW method for both software versions ($r= 0.79$ and $r=0.74$, for the v.5.1 and v. 6.1, respectively). This study also reported a similar type of systematic bias, in which the error was dependent on physical activity level. In the current study, the larger differences in TEE estimation when comparing the same software version (6.1) utilized in Arvidsson's study (et al. 2009b) can be attributable to a larger age range, with a probable larger maturation range. In concordance with Arvidsson's study, we also found significant improvements in TEE estimation with the more recently developed algorithms, high correlation coefficient for both monitors and algorithms, and a similar form of systematic bias observed with increased intensities. The systematic bias reported by Arvidsson (et al. 2009b) and observed in the current study might suggest a limitation of the SWA monitor to properly assess EE at higher intensities, as recently reported in two studies including highly trained athletes (Drenowatz et al. 2010; Koehler et al. 2010).

The continued release of newly developed SWA software makes it difficult to compare results across studies; however, it seems reasonable to assume that the latest software version maintains the positive characteristics observed in preceding studies, while incorporating additional capabilities to the device. The results in our study comparing previous algorithms (v. 2.2) with newly developed algorithms (v. 5.0) appears to show substantial improvement for assessing TEE in youth under free-living conditions.

Conversely, in another recent study with children (Bäcklund et al. 2010), researchers assessed the validity of the SWA under free-living conditions, using the DLW method in a sample of obese and overweight children. In the study, a sample of 22 healthy, overweight, or obese children (aged 8-11 yrs.) were recruited to simultaneously assess their TEE during 14-days of monitoring via DLW and SWA estimation. The researchers also used two different versions of the SWA software (V. 5.1 and V. 6.1) to process the results and compare their validity. Surprisingly, the results showed non-significant differences (<1% difference) in group level TEE with the SWA when using the software version 5.1, but a significant difference (18% underestimation) when using the software version 6.1. The results from that study contradict the findings from Arvidsson and colleagues (2009b), perhaps suggesting a better estimation of TEE in overweight and obese children with adult SWA algorithms compared to the more recently developed children algorithms. The discrepancies in TEE estimation between different SWA studies involving children could be explained by the different versions of the software utilized, the intensity of the activities included in the study, or the specific characteristics of the sample involved in the study (i.e., BMI category, maturation status).

While estimates of TEE were reasonable in the present study, we observed considerable error in the estimates of AEE (despite the presence of high correlations). Activity energy expenditure is the most variable component of total daily energy expenditure (Levine et al. 2004) and includes a wide variety of activities of daily living. The diverse and variable nature of lifestyles makes it very difficult to accurately assess AEE. However, the results with the SWA are better than reported with other monitors in

similar DLW studies. A study by Nilsson et al. (2007) showed large variation in estimation (ranging from 83% overestimation to 46% underestimation) dependent on the type of equation utilized, and how the equation was developed (i.e.: sample utilized, laboratory conditions vs. free-living). A study by Ekelund and colleagues (2001) showed that gender and physical characteristic variables (height, weight, fat-free mass) contribute to the variability and error in estimates of AEE. Additional work is needed to understand this error but the integration of multi-sensors and the use of complex pattern recognition technology in the SWA platform appear to offer advantages over other monitoring technology.

One of the strengths of the study is the high compliance displayed by the participants. Large compliance gaps during the protocol could introduce larger error when making assumptions about the intensities of the missing minutes. In Arvidsson's and colleagues (2009a) previous free-living study with the SWA in youth, researchers reported similar compliance values ($97.2 \pm 2.3\%$) to our study ($96.7 \pm 3.0\%$). The utilization of a convenience sample for the study is a limitation of the current study. The characteristics of the study require responsible and reliable participants; therefore, as previously stated, internal validity was weighed more heavily than external validity. As a result, the findings in the current study cannot be generalized.

In conclusion, the newly developed SenseWear 5.0 algorithms outperformed the version 2.2 algorithms. The low absolute error rates of the SWA and Mini monitor support the validity of the new algorithms for assessing TEE under free-living conditions

in youth. Additional research on the SWA for use in youth populations should focus in understanding factors contributing to large individual variability.

Acknowledgements

The authors gratefully acknowledge the enthusiastic support of the volunteers who participated in this study.

Disclosures

This research was funded by a grant from BodyMedia Inc. awarded to Dr. Greg Welk.

Table 2. Descriptive statistics.

		Age (years)	Weight (Kg)	Height (cm)	BMI (kg/ m²)	Age PHV (years)
Boys N=15	Mean	12.9	52.9	161.7	19.6	13.3
	SD	2.1	19.1	16.3	4.1	0.7
	Range	10.2 - 16.3	26.2 - 91.4	129.0 - 185.6	14.1 - 26.9	11.9 - 14.2
Girls N=13	Mean	12.3	49.3	160.2	19.2	11.8
	SD	1	8.2	9.8	2.3	0.2
	Range	11.0 – 14.0	35.8 - 65.3	141.8 - 173.2	16.2 - 23.5	11.3 - 12.1
All N=28	Mean	12.4	51.3	161	19.4	12.6
	SD	1.6	14.9	13.5	3.3	0.9
	Range	10.2 - 16.3	26.2 - 91.4	129.0 - 185.6	14.1-26.9	11.3 - 14.2

Table 3. Total energy expenditure (TEE) values from the activity monitors and DLW method.

A. Currently available algorithms (kcal·day⁻¹).

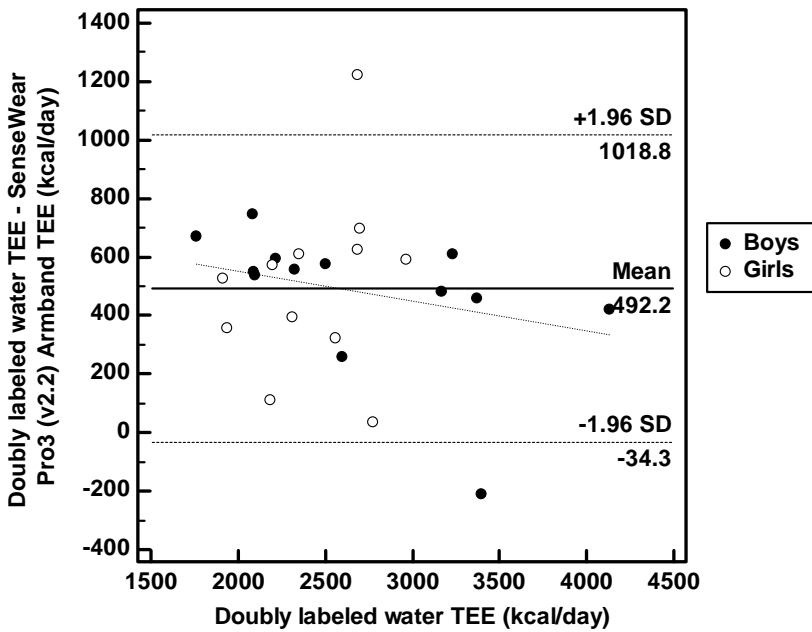
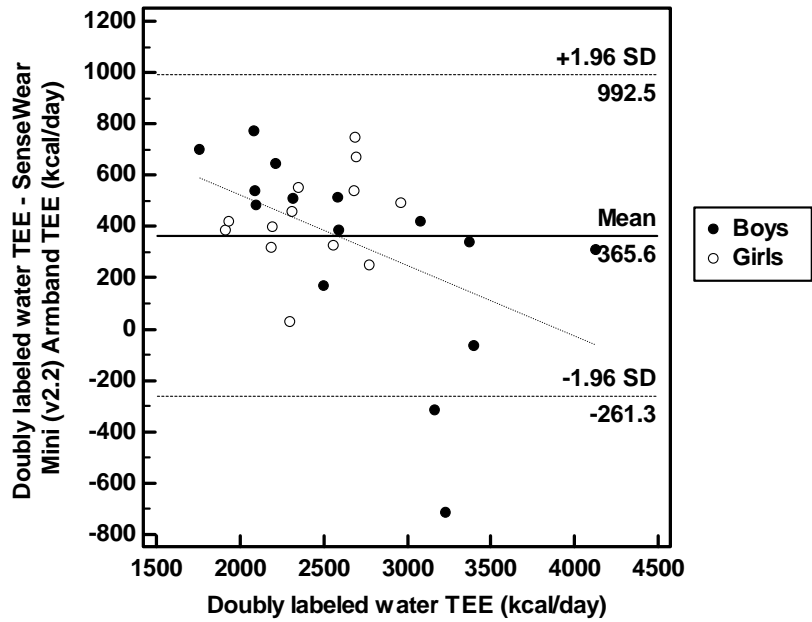
		TEE _{DLW}	TEE _{Mini 2.2}	TEE _{SWA 2.2}
Boys N=15	Mean	2690.2	2402.4	2209.8
	SD	701.6	1020.5	847.5
	Range	1760.0- 4132.0	1062.3- 3949.1	1089.8- 3712.6
Girls N=13	Mean	2429.1	2000.9	1934.6
	SD	326.7	318.8	398.5
	Range	1915.0- 2963.0	1522.0-2526.4	1388.1- 2739.0
All N=28	Mean	2579.3	2213.7	2077.7
	SD	540.6	745.0	672.0
	Range	1760.0- 4132.0	1062.3- 3949.1	1089.8- 3712.6

B. Newly developed algorithms (kcal·day⁻¹).

		TEE _{DLW}	TEE _{Mini 5.0}	TEE _{SWA 5.0}
Boys N=15	Mean	2690.2	3047.5	2928.2
	SD	701.6	1052.4	983.4
	Range	1760.0- 4132.0	1640.6- 4628.6	1596.0- 4534.9
Girls N=13	Mean	2429.1	2526	2438.9
	SD	326.7	398.6	496.9
	Range	1915.0- 2963.0	1804.0- 3378.9	1776.6- 3451.4
All (N=28)	Mean	2579.3	2785.3	2693.3
	SD	540.6	801.5	811.7
	Range	1760.0- 4132.0	1640.6- 4628.6	1596.0- 4534.9

Figures

Figure 1- Bland-Altman plots for total energy expenditure (kcal·day⁻¹).



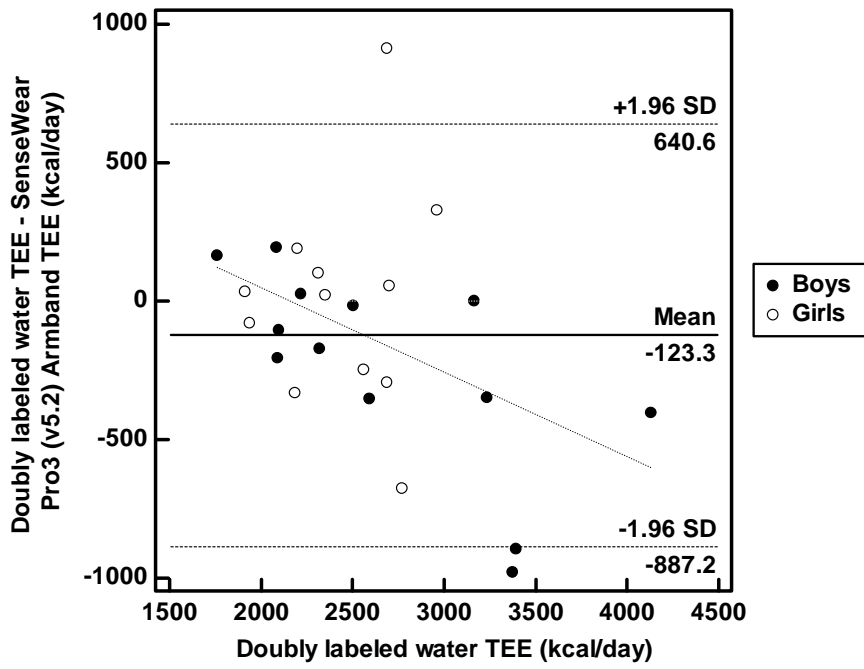
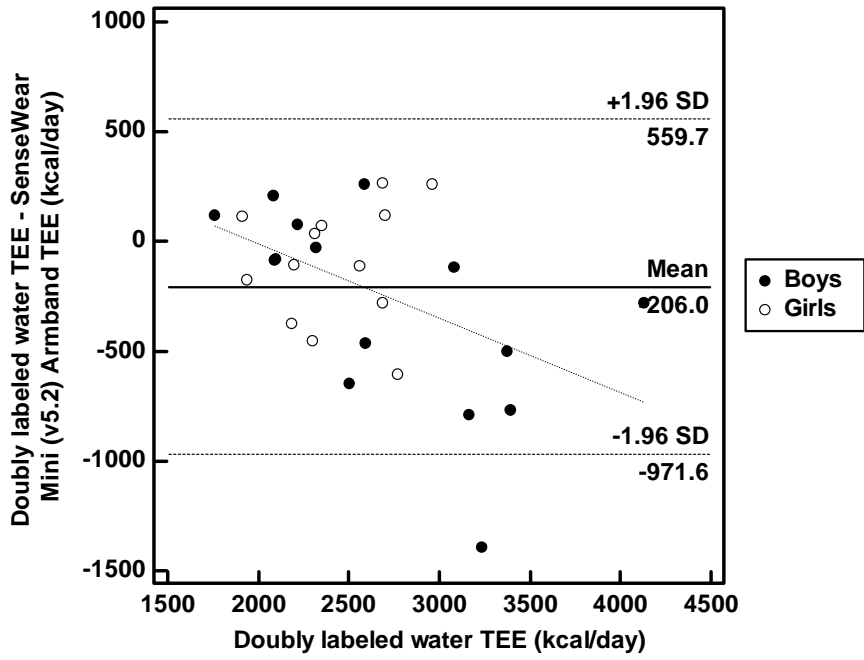
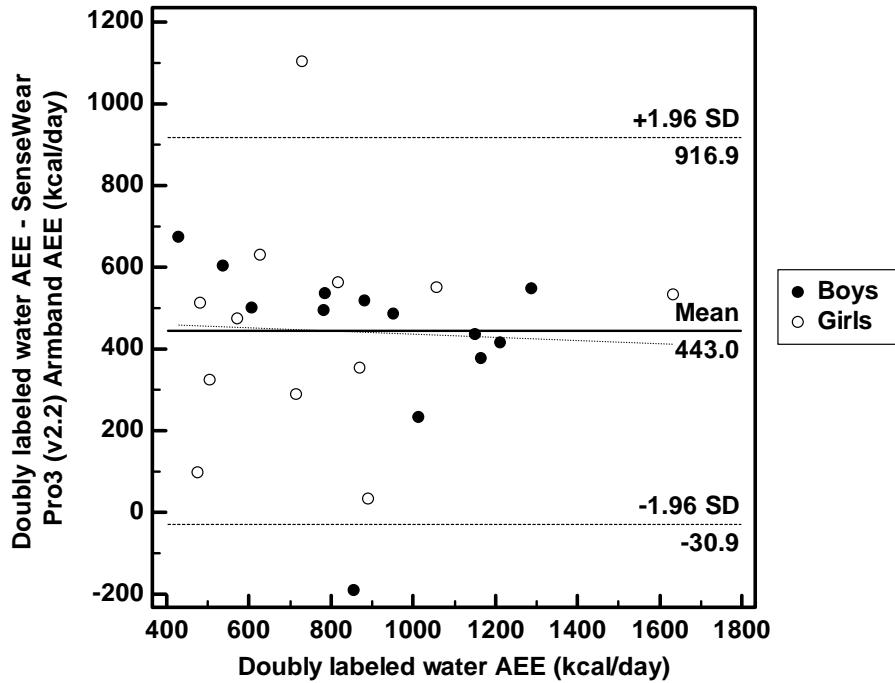
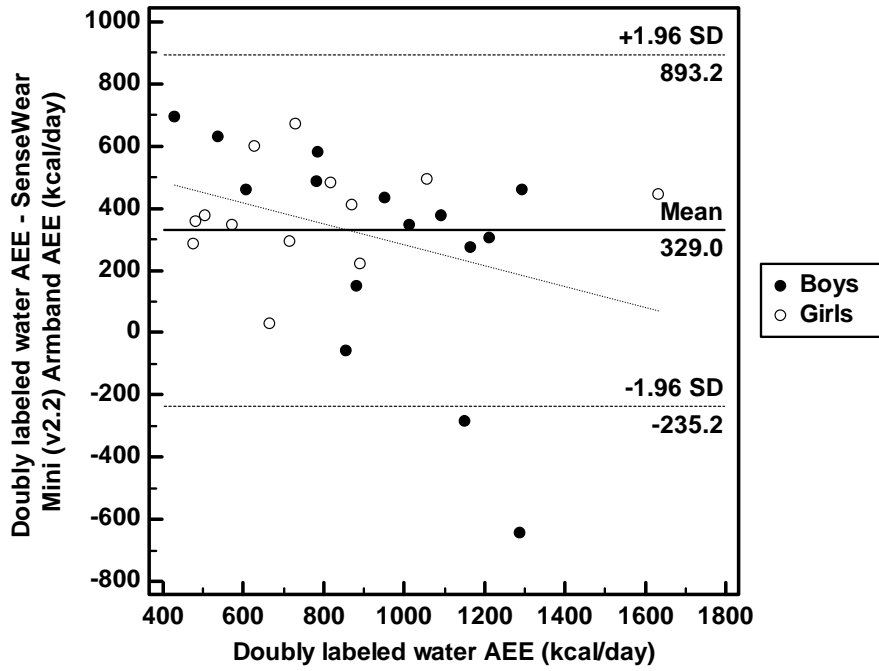
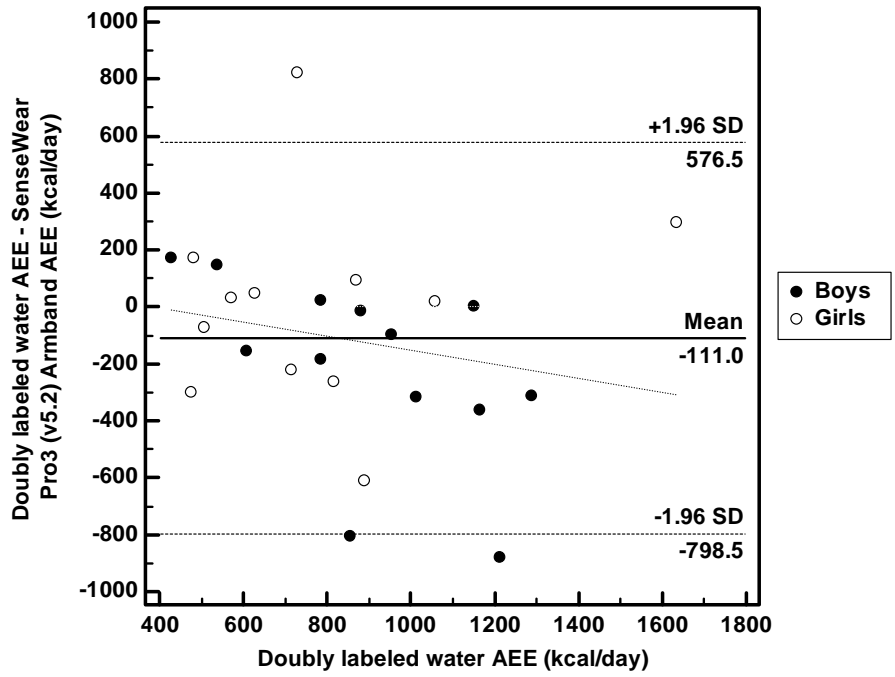
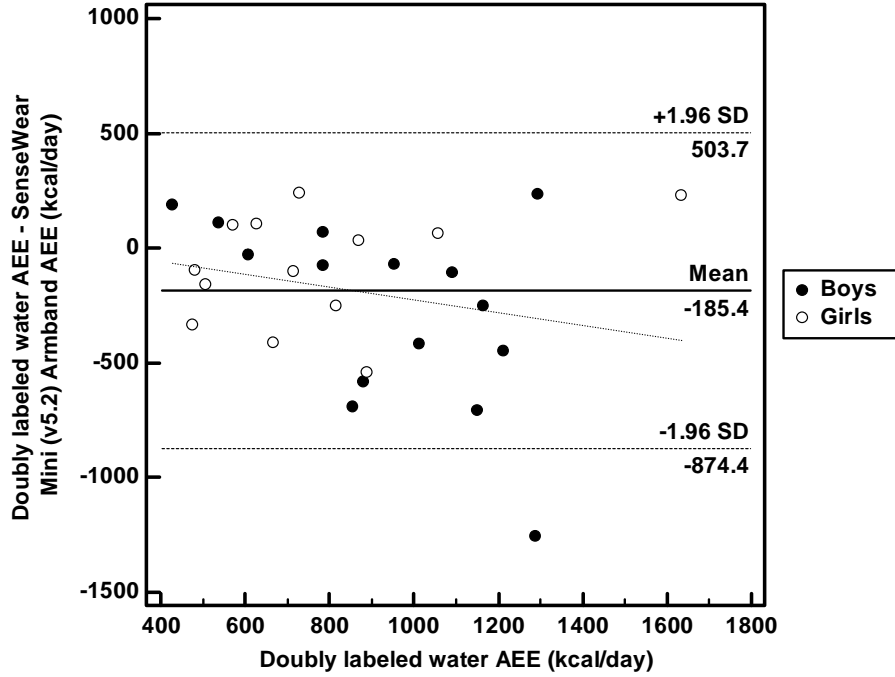


Figure 2- Bland-Altman plots for activity energy expenditure (kcal·day⁻¹).





References

Arvidsson D, Slinde F, Larsson S, Hulthen L. Energy cost of physical activities in children: validation of SenseWear Armband. *Med Sci Sports Exerc.* 2007;39(11):2076-2084.

a-Arvidsson D, Slinde F, Larsson S, Hulthén L. Energy cost in children assessed by multisensor activity monitors. *Med Sci Sports Exerc.* 2009;41(3):603-11.

b-Arvidsson D, Slinde F, Hulthen L. Free-living energy expenditure in children using multi-sensor activity monitors. *Clinical Nutr.* 2009; 28:305-312.

Bäcklund C, Sundelin G, Larsson C. Validity of armband measuring energy expenditure in overweight and obese children. *Med Sci Sports Exerc;* 2010;42(6):1154-61.

Bland JM and Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* . 1986; 1: 307-310.

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

Corder K, Brage S, Mattocks C, et al. Comparison of two methods to assess PAEE during six activities in children. *Med Sci Sports Exerc.* 2007;39(12):2180–8.

Drenowatz C, Eisenmann JC. Validation of the SenseWear Armband at high intensity exercise. *Eur J Appl Physiol*. 2010 Oct 24. [Epub ahead of print]

Dorminy CA, Choi L, Akohoue SA, Chen KY, Buchowski MS. Validity of a multisensor armband in estimating 24-h energy expenditure in children. *Med Sci Sports Exerc*. 2008; 40(4):699-706.

Ekelund U, Sjöström M, Yngve A, Poortvliet E, Nilsson A, Froberg K, Wedderkopp N, Westerterp K. Physical activity assessed by activity monitor and doubly labeled water in children. *Med Sci Sports Exerc*. 2001 Feb;33(2):275-81.

Harrell JS, McMurray RG, Baggett CD, Pennell ML, Pearce PF, Bangdiwala SI. Energy costs of physical activities in children and adolescents. *Med Sci Sports Exerc*. 2005 Feb;37(2):329-36.

Johannsen DL, Calabro MA, Stewart J, Franke W, Rood JC, Welk GJ. Accuracy of armband monitors for measuring daily energy expenditure in healthy adults. *Med Sci Sports Exerc*. 2010; 42(11):2134-40.

Koehler K, Braun H, de Marées M, Fusch G, Fusch C, Schaenzer W. Assessing energy expenditure in male endurance athletes: validity of the Sensewear Armband. *Med Sci Sports Exerc.* 2010 Dec 1. [Epub ahead of print]

Levine JA. Nonexercise activity thermogenesis (NEAT): environment and biology. *Am J Physiol Endocrinol Metab.* 2004; 286; E675-E685.

Mirwald RL, Baxter-Jones ADG, Bailey DA, Beunen GP. An assessment of maturity from anthropometric measurements. *Med Sci Sports Exerc.* 2002; 34(4); 689-694.

Nilsson A, Brage S, Riddoch C, Anderssen SA, Sardinha LB, Wedderkopp N, Andersen LB, Ekelund U. Comparison of equations for predicting energy expenditure from accelerometer counts in children. *Scand J Med Sci Sports.* 2008 Oct;18(5):643-50.

Ridley K, Ainsworth BE, Olds TS. Development of a Compendium of energy expenditures for youth. *Int J Behav Nutr Phys Act.* 2008; 5: 45.

Slootmaker SM, Schuit AJ, Chinapaw MJM, Seidell JC, van Mechelen W. Disagreement in physical activity assessed by accelerometer and self-report in subgroups of age, gender, education and weight status. *Int J Behav Nutr Phys Act.* 2009; 6: 17.

St-Onge M, Mignault D, Allison DB, Rabasa-Lhoret R. Evaluation of a portable device to measure daily energy expenditure in free-living adults. *Am J Clin Nutr* 2007;85(3):742-749.

Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. *Med Sci Sports Exerc.* 2005;37(11 Suppl):S531-543.

Welk GJ, McClain JJ, Eisenmann JC, Wickel EE. Field validation of the MTI Actigraph and BodyMedia armband monitor using the IDEEA monitor. *Obesity* (Silver Spring). 2007 Apr;15(4):918-28.

CHAPTER 5. SUMMARY

An extensive body of literature has documented the health benefits of physical activity (Lavie et al. 2011). However, studies have shown that the relationship between PA behavior and health outcomes is not as strong as the relationship between cardiorespiratory fitness and those same health outcomes (Williams 2001). One reason why the relationship is not as strong is due to the inherent challenges in assessing physical activity behavior (Lee et al. 2010) compared to cardiorespiratory fitness.

The assessment of PA under free-living conditions has proven to be particularly challenging. The “gold standard” for energy expenditure assessment under free-living conditions, the doubly labeled method, has shown to have a 5.1% measurement error from a combination of analytical variation (2.9%) and physiologic variation (4.2%) (Trabulsi et al. 2003). The procedure is somewhat “cumbersome” and very costly, which makes it unfeasible for large scale studies. Self-report instruments are inexpensive, easy to administer and allow for characterization of activities, domains, and purpose of activity. However, self-report instruments generally provide only rough estimates and have limited utility for many research applications. Accelerometry-based activity monitors provide a reasonable compromise in terms of validity and feasibility and have become the de-facto standard for contemporary physical activity research. These devices, while widely used, still have significant limitations for assessing free-living physical activity behavior. They are designed to capture locomotor activity but the majority of a person’s day is spent in lower intensity (Rest or Light) activities. The limitations of standard accelerometry-based monitors for assessing activities of daily living has been

clearly documented (Crouter et al. 2006) and this has led to interest in alternative measurement approaches. New pattern-recognition techniques, such as artificial neural networking and Hidden Markov Modeling (Poher et al. 2006; Staudenmayer et al. 2009), have been recently developed to improve the accuracy of standard monitors.

The BodyMedia Armband is unique since it incorporates multiple sensors and features built-in pattern recognition technology. A growing body of literature has supported the validity of the BodyMedia monitors under different conditions and with different samples (See literature review). However, additional validation research was needed to further evaluate these monitors, particularly in young children and older adults. These are populations that present unique challenges for the assessment of physical activity. Most previous studies were conducted under laboratory conditions, so additional work was also needed to evaluate validity under free-living conditions. The overarching theme of my dissertation research has been on addressing these gaps in the research and contributing new insights into the validity of the Sensewear monitors for evaluating PA and EE

A number of preliminary studies used indirect calorimetry as the criterion measure but my dissertation research used the strongest possible criterion measure to evaluate the monitors under free-living conditions. In a previous study in our laboratory (Johannsen et al. 2010), we tested the accuracy of two armband monitors (SenseWear Pro₃ Armband and SenseWear Mini Armband) under free-living conditions in young adults. Low mean absolute error rates (~8%) supported the validity of the monitors to assess total energy expenditure under free-living conditions. The results of the present

studies (in the table below) have extended this work by demonstrating similarly low absolute error rates for assessments in children and older adults.

	Age range (years)	TEE (MAPE)	AEE (MAPE)	Study
Children	10-16	10.4-11.4	31.5-32.5	Study II
Adults	24-60	8.1-8.3	26.0-28.0	Johannsen et al. 2010
Older adults	60-78	6.5-8.4	17.1-21.9	Study I

While 6.5-11.4% mean absolute percentage error (MAPE) for TEE may seem high, these values are considerably better than other previously reported accelerometry-based activity monitors in DLW studies (-29 to 24% error; Leenders et al. 2006; Plasqui et al. 2007). Collectively, the findings from this series of studies provide strong support the validity of the SenseWear Armband for assessing physical activity and energy expenditure under free-living conditions in children, young adults and older adults.

Pattern-recognition activity monitors continue to show advances in PA estimation under free-living conditions. The advances continue to narrow the error gap between the PA estimators and the criterion values. Additional advances in technology and assessment research will make it possible to assess physical activity more accurately and this will help to advance physical activity research in a number of areas.

References

Crouter SE, Churilla JR, Bassett DR Jr. Estimating energy expenditure using accelerometers. *Eur J Appl Physiol* 2006;98(6):601-612.

Johannsen DL, Calabro MA, Stewart J, Franke W, Rood JC, Welk GJ. Accuracy of armband monitors for measuring daily energy expenditure in healthy adults. *Med Sci Sports Exerc.* 2010; 42(11):2134-40.

Lavie CJ, Church TS, Milani RV, Earnest CP. Impact of physical activity, cardiorespiratory fitness, and exercise training on markers of inflammation. *J Cardiopulm Rehabil Prev.* 2011 Mar 21. [Epub ahead of print]

Lee DC, Sui X, Ortega FB, Kim YS, Church TS, Winett RA, Ekelund U, Katzmarzyk PT, Blair SN. Comparisons of leisure-time physical activity and cardiorespiratory fitness as predictors of all-cause mortality in men and women. *Br J Sports Med.* 2010 Apr 23. [Epub ahead of print]

Leenders NY, Sherman WM, Nagaraja HN. Energy expenditure estimated by accelerometry and doubly labeled water: do they agree? *Med Sci Sports Exerc.* 2006 Dec;38(12):2165-72.

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

Pober, D.M., Staudenmayer, J., Raphael, C., Freedson, P.S. Development of novel techniques to classify physical activity mode using accelerometers. *Med. Sci. Sports Exerc.* 2006; 38(9):1626-34.

Staudenmayer, J., Pober, D., Crouter, S.E., Bassett, D.R., and Freedson, P. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. *J Appl Physiol.* 2009; 107(4):1300-7.

Trabulsi J, Troiano RP, Subar AF, Sharbaugh C, Kipnis V, Schatzkin A, Schoeller DA. Precision of the doubly labeled water method in a large-scale application: evaluation of a streamlined-dosing protocol in the Observing Protein and Energy Nutrition (OPEN) study. *Eur J Clin Nutr.* 2003; 57(11):1370-7.

Williams, P. T. Physical fitness and activity as separate heart disease risk factors: a meta-analysis. *Med. Sci. Sports Exerc.* 2001; 33(5):754-761.

APPENDIX A

Validation of the SenseWear Pro Armband algorithms in children

Article published in Medicine and Science of Sports and Exercise.

Authors:

Calabró, Miguel A.¹, Welk, Gregory J.¹, Joey C. Eisenmann².

¹ Department of Kinesiology, Iowa State University, Ames, IA.

² Department of Kinesiology, Michigan State University, East Lansing, MI.

Reference:

Calabró MA, Welk GJ, Eisenmann JC. Validation of the SenseWear Pro Armband algorithms in children. *Medicine & Science in Sports & Exercise*. 2009 Sep; 41(9):1714-20.

ABSTRACT

Introduction: The SenseWear Pro Armband (SWA) has been shown to be a valid and practical tool to assess energy expenditure in adults. Recent studies have reported significant errors in energy expenditure estimates when the algorithms are applied to children. The purpose of this study was to assess the validity of new proprietary algorithms that were developed to take into account children's unique movement patterns. **Methods:** Twenty one healthy children (14 boys, 7 girls), averaging 9.4 (± 1.3) years of age, participated in a range of activities while being monitored with the SWA and a metabolic analyzer (TrueMax 2400, ParvoMedics, UT, USA). The activity protocol lasted 41 minutes and included: resting (lying down), coloring (sitting), playing computer games, walking on a treadmill (2, 2.5 and 3mph) and stationary bicycling. **Results:** The original algorithms overestimated energy expenditure (EE) by 32%, but average error with the newly developed algorithm was only 1.7%. There were no significant differences in overall estimates of EE across the 41 minute trial ($p > .05$) but there was some variability in agreement for specific activities (average absolute difference in EE estimates was 13%). The average error in EE estimates with the new algorithms were -20.7, -4.0, -4.9, -0.9, 0.6, 3.5 and -25.1 % for resting, coloring, computer games, walking on a treadmill (2, 2.5 and 3mph) and biking, respectively. Biking was the activities with significant differences in EE estimations ($p < 0.001$). Average minute by minute correlations across individuals was $r = 0.71 \pm 1.3$ indicating that the relationships were consistent across individuals. **Discussion:** The newly developed algorithms demonstrate improved accuracy for assessing energy expenditure in children – including accurate

estimation of light activities. **Conclusions:** The newly developed SWA algorithms provide accurate estimates of EE for typical activities in children.

INTRODUCTION

Physical activity is a complex behavior and it has proven to be very difficult to assess in field based research. The development of valid and reliable measurement tools is essential for advancing research on physical activity. Accelerometry-based activity monitors have become the most widely strategy for assessing physical activity under free living conditions but, despite considerable work, many challenges remain (Troiano, 2005). Researchers interested in youth behavior have additional challenges to overcome including more sporadic physical activity patterns in children (Freedson, 2005) and inherent variability due to growth and maturation (Wickel, 2008).

Pattern-recognition activity monitors offer considerable promise for improving the accuracy of physical activity assessment techniques. The Sensewear armband (SWA) monitor, for example, integrates motion sensor data with physiological data to estimate the energy cost of free living activity. An advantage of this multi-channel approach is that the heat related sensors provide additional information that can't be obtained solely from movement sensors. The heat related sensors, for example, provide a way to assess the energy cost of complex, non-ambulatory activities. The sensors can also detect the increased work required to walk up a grade or to carry a load (McClain, 2005). The validity of energy expenditure (EE) estimates from the SenseWear Pro Armband (SWA) has been supported in studies using both indirect calorimetry (Jakicic 2004; Fruin 2004; King 2004) and doubly labeled water (Mignault 2005; St-Onge 2007). Recent research has demonstrated potential advantages of the SWA when compared with traditional accelerometry-based monitors (Welk, 2007).

Results of validation studies in youth have been more equivocal. Arvidsson and colleagues (2007) reported that the SWA significantly underestimated EE for a variety of standardized physical activities in a sample of 20 children. In contrast, a study by Dorminy et al. (2008) reported consistent overestimation of EE with the SWA in a sample of 21 youth. The nature of the discrepancies in these results is not clear but it is not completely surprising since the propriety SWA algorithms were developed primarily on adults. The purpose of this study was to assess the validity of new proprietary algorithms that were developed specifically from children's data. Comparisons are made between estimates from the currently available algorithms and new algorithms (provided by the manufacturer) to clarify the nature of the errors reported in previous studies.

METHODS

Participants

Twenty two healthy children were recruited from a summer youth fitness camp hosted by the local University. The camp provides activity programming to youth in the summer (as a form of day care) and tends to attract participants from diverse cultural backgrounds (28 % minority) and socio-economic backgrounds. While it is an activity-based program, the participants are not particularly athletic so there is also diversity with regard to activity level. Approval from the Institutional Review Board was obtained before the beginning of the study. Parental consent and children's assent were obtained after informing about the procedures and purposes of the study. One of the participants had to be excluded from the analysis due to faulty metabolic analyzer data.

Description of the SenseWear Pro 2 Armband (SWA)

The SWA is a wireless, non-invasive, multi-sensor activity monitor that is worn over the triceps muscle. The SWA armband monitor integrates data from 5 sensors including a 2-axis accelerometer, heat flux sensor, galvanic skin response (GSR) sensor, skin temperature sensor, near body ambient temperature sensor to estimate energy expenditure under free living conditions. The heat related sensors provide additional information about the energy cost of activity since periods of increased work are associated with increased heat production. The GSR sensor may also contribute to EE estimation since it detects changes in the skin's electrical properties due to sweat gland activity and psychological stimulus (periods of increased stimulus are associated with increased skin conductance). The direct contributions of heat indices and GSR in the prediction algorithms are not shared by the company but all five channels are used in estimations of EE (BodyMedia, personal communication).

A unique aspect of the SWA monitor is that the company continues to upgrade and enhance the software as new training data become integrated into the pattern recognition algorithms. The manufacturer has recently collected data from three independent research teams to improve the accuracy of algorithms for children but they have yet to be formally released. The present study compared the estimates obtained directly from the current InnerviewTM Research Software (version 4.2) with new proprietary algorithms scheduled for next upgrade.

Data Collection

Participants were guided to the laboratory in their scheduled day of testing and were instructed about the characteristics of the study before signing assent documents. Anthropometric measures were obtained at the beginning of the data collection session.

Standing height was measured to the nearest 0.1 cm with the use of a wall mounted Harpenden stadiometer (Harpenden, London, UK) and with the participants barefooted. Body mass was measured with participants in light clothes and barefooted on an electronic scale (Seca 770) to the nearest 0.1 kg. The body mass index (BMI) was calculated as $\text{weight (kg)/height}^2 \text{ (m}^2\text{)}$.

After completing the anthropometric measurements, participants were fitted with the SWA monitor and a pediatric mask for use with the metabolic cart (TrueMax 2400 (ParvoMedics, UT, USA). The participants were asked to complete a 41 minute activity protocol designed to simulate a variety of typical activities for children. The protocol consisted of 7 activity stages (5-minutes each) separated by 1-minute resting intervals. Descriptions of each activity are provided below:

Resting: Participant's rested on a lab table, in a supine position, during the 5-minutes of the stage. Lights were kept on and participants were instructed not to talk during the stage.

Coloring: Participant's selected animal drawings from a group of different figures and proceeded to color them with crayons. The speed of coloring was selected by the participant and the activity was finished after five minutes, even without completion of the task.

Computer games: Participants engaged in computer games on a desktop personal computer throughout the duration of the stage. The game selected involved pushing keys on the keyboard and did not include the use of a joysticks or a similar device.

Walking paces (3): Three walking paces of 2.0, 2.5 and 3.0 mph were completed by the participants on a treadmill. The participants were instructed not to use the treadmill handrails.

Biking: Sit height was adjusted to participant's leg length. Participants were instructed to pedal at 60 rpm with 0.5 Kp of resistance.

Data Processing

Breath by breath data from the metabolic cart were downloaded and aggregated to provide average minute by minute values to facilitate integration with the SWA data. The SWA armband data were downloaded using the Innerview Research Software (version 4.2). The raw output (.swd) file from the software was also sent directly to the company to obtain the estimates with the revised algorithms. The company provided a corresponding minute-by-minute estimate with the newly developed algorithm and these were merged with the metabolic data and the data obtained directly from the software.

Data Analyses

Traditional measurement agreement analyses were used to evaluate the validity of the two SWA algorithms. The primary statistical analyses involved evaluating overall group differences in EE estimates from the three methods across the whole monitoring period (41 minute trial). Many validation studies have focused comparisons on evaluations of point estimates of individual physical activities but consideration also needs to be given to the overall accuracy during a sustained period of monitoring. Additional analyses were conducted to evaluate the accuracy of point estimates for each stage in order to determine how errors in individual activities impact the overall estimates. Mixed model analyses of variance were used to account for the possible

correlation across repeated observations taken on the same individuals in the study. The models (run in SAS 9.0) used participant within gender as person-level random effect term and the residual variance as a second random effect term. These analyses assume a common variance for among-person and within-person random effects. The fixed effects included in the models for EE were Gender and Method (IC, SWA, SWAold). F-tests were used to determine if factors were statistically significant and Tukey-Kramer paired comparisons tests were used to test for differences among levels of fixed effects. Least squares means and standard errors for all effects were estimated within the model, and these values are reported in the descriptive tables.

Additional analyses were conducted to evaluate overall measurement agreement. Minute by minute correlations were computed for each participant's set of data and the mean correlation coefficient across participants was used to reflect the overall association. Consistent with contemporary measurement research, we also utilized Bland Altman graphical procedures (Bland 1986) to examine agreement across the range of intensities. The mean of the two estimates (x-axis) is plotted against the difference between the two estimates (y-axis) to allow for detection of systematic forms of bias in the estimates. Confidence intervals defining the limits of agreement were established as 1.96 SD from the mean difference.

RESULTS

The study evaluated the agreement between measured EE and two estimates of EE from the SWA monitor in a diverse sample of youth aged (7-12). While the sample was predominantly white (72%), the demographics were typical of the surrounding

community and more diverse than the state as a whole. There was a range of body types with approximately 14% characterized as “at risk for overweight (between 85th and 95th percentile) and 5% characterized as “overweight” (> 95th percentile). Descriptive statistics for the sample population are provided in Table 1.

The primary analyses involved method comparisons of the overall EE estimates. Mixed model analyses revealed significant method effects [F (2, 2393)= 66.81, p<0.001]. Post hoc tests revealed that there were significant differences between the EE estimate from IC and the EE estimate from the existing SWA algorithm (F = -9.68, p < 0.0001). Least square mean differences revealed that the old algorithm differed from the IC value by an average of over .5 kcal/min (0.52 ± 0.05). Expressed differently, this equates to an average overestimation of approximately 32%.

Mixed model analyses also revealed significant differences between the estimates from the old SWA algorithm and the new algorithm (F = -10.31, p < 0.0001). The new values were significantly lower and this correction led to non-significant differences between the new SWA estimates and the criterion measure (F = 0.63, p = 0.53). Least square mean differences between IC and the new SWA values varied by only 0.03 ± 0.05 kcal/min (error of approximately 1.7%).

Subsequent analyses examined the agreement for individual activities performed during the 41 minute protocol. The average absolute difference in EE estimates between methods for the various activities was 13% with the new algorithm. The stage-specific EE differences were -0.25, -0.05, -0.05, -0.02, 0.02, 0.09 and -1.00 (in kcal/min) for resting, coloring, computer games, walking on a treadmill (2, 2.5 and 3mph) and biking, respectively. Biking was the only activity where EE estimations were significantly

different ($p < 0.001$). The data were processed separately by gender and the relationships were similar for both males and females (data not reported).

To evaluate overall measurement agreement, we computed minute by minute correlations across the 41 minute-trial for each participant (see Table 3). The average minute-by-minute correlation ranged from 0.36 to 0.87 with the new algorithm (mean = 0.71). The values for the old algorithm exhibited a similar range (from 0.46 to 0.96) and average (mean = 0.72) suggesting that the revised algorithms shifted the estimates in a consistent way across individuals. Figure 1 shows a plot of the average minute by minute correlation for both the old and new algorithm compared with the measured EE. The plot shows that both algorithms track the overall pattern of EE but the estimates from the new algorithm exhibited lower error and improved fit. Examination of the individual values showed no systematic differences in the magnitude of the correlations across participants.

The Bland Altman plots in Figure 2 provide a more detailed view of the differences in measurement agreement between the measures. The plot shows a tighter clustering of data points about the mean for the new algorithm and less overall error compared with the measured EE values. A cluster of points at the upper right of the plot show the continued underestimation with the estimate of biking activities. There was no evidence of any systematic bias across the range of EE values measured in the study.

DISCUSSION

This study examined the agreement in EE estimates from the SenseWear Pro Armband monitor in children. A unique aspect of the study is that we directly compared the existing algorithm with a newly developed version that will be released in a

subsequent firmware upgrade released to all users. The results demonstrated the new algorithms yield more accurate estimates than the existing version (4.2) that are currently available within the software.

Two previous studies reported limitations with the use of the SWA in children but the results were inconsistent. Work by Arvidsson et al (2007) showed a tendency for underestimation, while work by Dorminy et al. (2008) reveal a tendency for overestimation. The nature of the sample and the specific selection of activities can influence the estimates of EE in these types of studies. Variability in metabolic carts can also contribute error. Our results are more consistent with the findings by Dorminy that indicate a tendency for the current SWA algorithms to overestimate EE in youth. We found average overestimations of approximately 32% across the 41 minute protocol. The effect was consistent across most of the activities and across individuals.

The manufacturer of the Sensewear Pro (Bodymedia Inc) developed the existing algorithms with a predominantly adult population and likely used extrapolations to create estimates for youth. The newly developed algorithms were created based on data obtained from three independent labs using slightly different protocols and testing different activities. The results from this study show that these new algorithms provide accurate estimates of EE for most activities. The only activity tested that had values significantly different than the cart values was biking. This activity has been notoriously difficult to assess with accelerometers and our data suggest that more work is needed to improve accuracy with the SWA armband. While the effect size for this difference in biking was still large (0.74), the error is lower than is typically reported with other monitors for biking activities.

The new algorithms assessed in this study resulted in improved EE estimation for a variety of sedentary, light and moderate-intensity activities. Arvidsson (et al. 2007) reported average error values of -18.6% for resting activities and -35.7% for light intensity game playing. The values in the present study show underestimation for resting activities (-20.7%) but low amounts of error for other light activities (-4.0% for coloring and -4.9 % for computer games). Dorminy (et al. 2008) reported overestimation of EE for resting activity (21.2%) and also for other sedentary activities (21.1%).

During walking, Arvidsson et al. (2007) reported average errors of 0.8% (1.9 mph), -8.6% (2.5 mph) and -9.7% (3.1 mph). In contrast, Dorminy et al (2008) reported an average error of 14.2% during walking at speeds ranging from 2.5 to 4.5 mph. In the current study, the average error for the different speeds were considerably lower. The average errors were -0.89%, 0.64% and 3.45% for 2.0, 2.5 and 3.0 mph walking, respectively. These results demonstrate a clear improvement in the accuracy of the SWA algorithms for walking and locomotor activities.

In general, researchers have found it more difficult to accurately capture the EE cost of activity in youth. A recent study by Trost and colleagues (2005) assessed the accuracy of different Actigraph equations developed to estimate EE in children. During walking at 3 and 4 mph, the average errors ranged from 13.3-23.3% (Trost equation), 23.5-29.4% (Freedson equation for children) and 0.6-13.3% (Puyau equation). A study by Corder and colleagues (2005), showed the Actigraph to overestimate EE during flat walking by 42.6% compared to indirect calorimetry. In the same study, the Actical monitor overestimated the energy cost of flat walking by 33.3%, while the Actiheart monitor (combines accelerometry with HR) overestimated by 5.6%. The values with the

new SWA algorithms yielded errors ranging from -0.89-3.45% for the three walking paces used in this study so the results compare favorably to the Actiheart estimates.

Accelerometry-based activity monitors have also been limited in their ability to assess the intensity of low intensity activities and lifestyle activities. A recent study by our group (Wickel et al, 2007) reported significant error when the Freedson equation is used to estimate the energy cost of low intensity activities. There are pattern recognition approaches in development that may improve the predictive accuracy of accelerometers (Poher; Crouter) but these have not been developed for use with children at this point. A recent study by Corder and colleagues (2007) evaluated the accuracy of eight different EE-prediction models to estimate EE for six different activities (2 sedentary and 4 non-sedentary) in children. The results revealed systematic errors for models incorporating accelerometry alone as well as for combination of accelerometry and HR. The systematic error was more pronounced in the accelerometry alone models showing that HR improved the accuracy – particularly for the lower intensity activities. Heart rate provides additional information to improve the prediction of EE, but the results from the present study demonstrate that the revised SWA algorithms can produce accurate EE estimates using a non-invasive armband that doesn't necessitate heart rate information. Heart rate information is typically not reliable across extended periods of monitoring due to the presence of artifact and missing data.

A novel aspect of the armband is that the multiple sensors may also enhance the accuracy of intermittent activities performed throughout the day. Most studies have focused on point estimates of specific activities but the plot shown in Figure 2 demonstrates that the SWA estimates closely mirror estimates during the transition

periods during activities. The ability of the SWA to estimate EE cost of light activities and to adjust to changes in the intensity of the activity is likely due to contributions from the heat sensors and GSR sensors. These sensors may pick up subtle changes that can't be inferred with accelerometers or heart rate information. Additional work is needed to better understand the contributions of the various data channels used in the SWA.

In conclusion, this study demonstrates that the SWA armband produces valid estimates of EE for assessing free living activities. The non-significant difference in EE across the full 41 minute protocol is also important since it provides an indicator of the monitor to assess EE across extended periods of time. A strength of the study is that we directly compared the previous equation to the newly developed one in order to directly determine the differences in estimates between these versions. A limitation is that we did not evaluate other higher intensity (vigorous) activities such as running. The goal of the study was to evaluate the accuracy of the monitor under real world conditions so emphasis was placed on selecting activities that were more reflective of a child's typical activity level. Most validation studies have focused on assessing specific physical activities but for energy balance research it is important for monitors to be able to assess a range of intensities including sedentary and light activities since these accounts for the bulk of the day. Additional work is clearly needed to examine the validity of the SWA armband across a wider range of activities, over longer periods of time and in different populations.

REFERENCES

- Arvidsson D, Slinde F, Larsson S, Hulthen L. Energy cost of physical activities in children: Validation of SenseWear Armband. *Med Sci Sports Exerc.* 39:2076-2084, 2007.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 8: 307-310, 1986.
- Corder K, Brage S, Wareham NJ, Ekelund U. Comparison of PAEE from combined and separate HR and movement models in children. *Med Sci Sports Exerc.* 37:2076-2084, 2005.
- Dorminy CA, Choi L, Akohoue SA, Chen KY, Buchowski MS. Validity of a multisensor Armband in estimating 24-h energy expenditure in children. *Med Sci Sports Exerc.* 40: 699-706, 2008.
- Freedson P, Pober D, Janz KF. Calibration of accelerometer output for children *Med Sci Sports Exerc.* 37:S523-S530, 2005.
- Fruin ML, Rankin JW. Validity of a multi-sensor armband in estimating rest and exercise energy expenditure. *Med Sci Sports Exerc.* 36:1063-1069, 2004.
- Jakicic JM, Marcus M, Gallagher KI, et al. Evaluation of the SenseWear Pro Armband to assess energy expenditure during exercise. *Med Sci Sports Exerc.* 36:897-904, 2004.
- King GA, Torres N, Potter C, Brooks TJ, Coleman KJ. Comparison of activity monitors to estimate energy cost of treadmill exercise. *Med Sci Sports Exerc.* 36:1244-1251, 2004.
- McClain JJ, Welk GJ, Wickel EE, Eisenmann JC. Accuracy of energy expenditure estimates from the Bodymedia Sensewear(R) Pro 2 Armband. *Med Sci Sports Exerc.* 37: S116-S117, 2005.

Mignault D, St Onge M, Karelis AD, Allison DB, Rabasa-Lhoret R. Evaluation of the portable HealthWear armband. *Diabetes Care*. 28:225-227, 2005.

St Onge M, Mignault D, Allison DB, Rabasa-Lhoret R. Evaluation of a portable device to measure daily energy expenditure in free-living adults. *Am J Clin Nutr*. 85:742-749, 2007.

Troiano RP. A timely meeting: objective measurement of physical activity. *Med Sci Sports Exerc*. 37:S487-S489, 2005.

Trost SG, Way R, Okely AD. Predictive validity of three Actigraph energy expenditure equations for children. *Med Sci Sports Exerc*. 38(2):380-387, 2005.

Welk GJ, McClain JJ, Eisenmann JC, Wickel EE. Field validation of the MTI Actigraph and BodyMedia Armband monitor using the IDEEA monitor. *Obesity*. 15:918-928, 2007.

Wickel EE, Eisenmann JC, Welk GJ. Predictive validity of an age-specific MET equation among youth of varying body size. *Eur J Appl Physiol*. 101:555-63, 2007.

Table 1. Sample characteristics (means \pm SD).

	<i>N</i>	Age (yrs)	Height (cm)	Weight (kg)	BMI (kg/m²)
Boys	14	9.4 \pm 1.4	137.2 \pm 9.9	32.7 \pm 6.4	17.2 \pm 1.6
Girls	7	8.4 \pm 2.3	124 \pm 34.8	29.2 \pm 9.3	18.0 \pm 4.0
All	21	9.4 \pm 1.2	136.2 \pm 8.5	32.6 \pm 6.7	17.5 \pm 2.6

Table 2. Stage energy expenditure (EE) values in children (N=21) (means \pm SD).

	IC	EE SWA	t-value	Pr>
Resting	1.18 \pm 0.31	0.94 \pm 0.16	1.91	0.057
Transition	1.00 \pm 0.33	0.92 \pm 0.40	0.33	0.742
Coloring	1.25 \pm 0.32	1.20 \pm 1.00	1.18	0.240
Transition	1.32 \pm 0.36	1.14 \pm 0.92	0.68	0.498
Computer games	1.13 \pm 0.44	1.07 \pm 0.40	0.38	0.705
Transition	1.27 \pm 0.34	1.16 \pm 0.49	0.44	0.662
Walking (2.0 mph)	2.25 \pm 0.61	2.23 \pm 0.80	0.29	0.773
Transition	1.77 \pm 0.37	2.15 \pm 1.07	-1.55	0.122
Walking (2.5 mph)	2.49 \pm 0.45	2.51 \pm 0.78	0.35	0.729
Transition	2.06 \pm 0.75	2.52 \pm 1.00	-1.85	0.065
Walking (3.0 mph)	2.67 \pm 0.55	2.77 \pm 0.73	-0.59	0.552
Transition	2.33 \pm 0.52	2.72 \pm 0.93	-1.47	0.143
Biking	3.99 \pm 1.35	2.99 \pm 0.82	7.68	<0.0001

IC= Energy expenditure measured with indirect calorimetry; EESWA= Energy expenditure estimated by the SenseWear Pro 2 Armband (SWA).

Table 3. Individual min-by-min correlations between the SWA estimates and IC estimates.

Participant	SWAnew	SWAold
2	0.87	0.88
3	0.36	0.57
4	0.68	0.93
5	0.72	0.68
6	0.75	0.66
7	0.78	0.62
8	0.60	0.64
9	0.87	0.89
10	0.59	0.81
11	0.84	0.96
12	0.67	0.93
13	0.71	0.73
14	0.51	0.83
15	0.74	0.72
16	0.76	0.46
17	0.80	0.71
18	0.74	0.46
19	0.53	0.58
20	0.84	0.56
21	0.75	0.76
22	0.79	0.79
Mean	0.71	0.72

Figure 1. Average minute-by-minute correlation for both the old and new SenseWear Armband (SWA) algorithm compared with the measured EE.

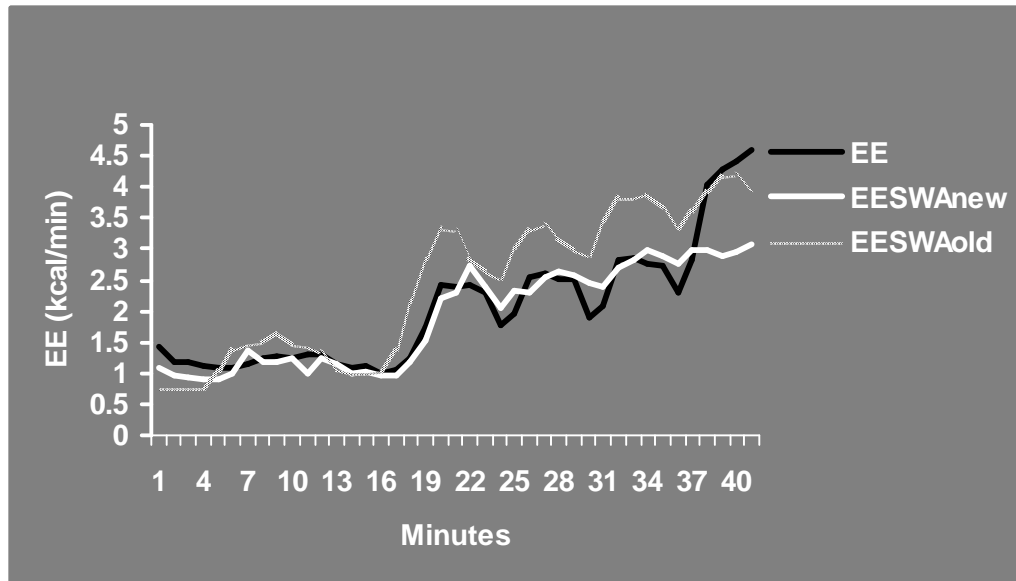
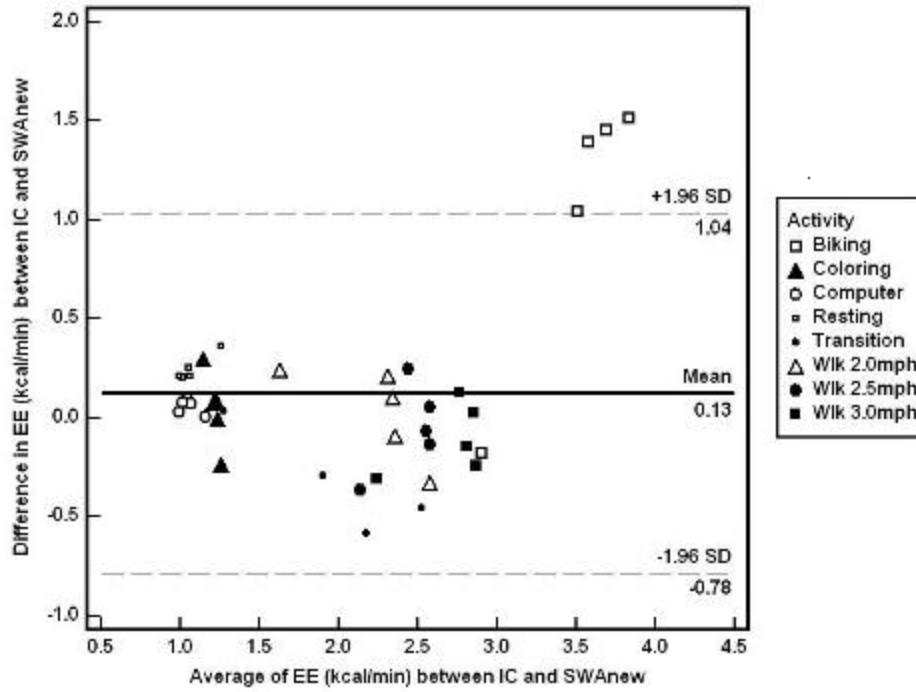


Figure 2. Comparison between the old and the new algorithms.



Accuracy of armband monitors for measuring daily energy expenditure in healthy adults

Article published in Medicine and Science of Sports and Exercise.

Authors:

Darcy L. Johannsen², M. Andres Calabro¹, Jeanne Stewart¹, Warren Franke¹, Jennifer C. Rood², Gregory J. Welk¹

¹ Iowa State University

² Pennington Biomedical Research Center

Reference:

Johannsen, DL, Calabro, MA, Stewart, J, Franke, W, Rood, J, and Welk, GJ. Validity of portable armband monitors for measuring daily energy expenditure compared with the doubly labeled water method in healthy adults. *Medicine and Science in Sports and Exercise*. 2010 Nov; 42(11):2134-40.

Abstract

Introduction: There is a need to develop accurate devices for measuring daily energy expenditure under free living conditions, particularly given our current obesity epidemic.

Purpose: The purpose of the present study was to evaluate the validity of energy expenditure estimates from two portable armband devices, the SenseWear Pro3 Armband monitor (SWA) and the SenseWear Mini armband monitor (Mini), under free-living conditions. **Methods:** Participants in the study (30 healthy adults aged 24-60 yr) wore both monitors for 14 consecutive days, including while sleeping. Criterion values for total energy expenditure (TEE) were determined using doubly labeled water (DLW), the established gold standard method for free-living energy expenditure assessment. **Results:** The average TEE estimates were within 112 kcal/day for the SWA and within 22 kcal/day for the Mini, but the absolute error rates (computed as average absolute value of the individual errors) were similar for the two monitors (SWA: $8.1 \pm 6.8\%$, Mini: $8.3 \pm 6.5\%$). Using intra-class correlation analysis, significant agreement was found between the SWA and DLW estimates of energy expenditure (ICC = 0.80, 95% CI: 0.89, 0.70) and between the Mini and DLW (ICC = 0.85, 95% CI: 0.92, 0.76). Graphical plots of the DLW TEE values against the difference between DLW and monitor estimates of TEE showed that agreement was consistent across a range of TEE values.

Key words: accelerometer, activity monitor, armband, doubly labeled water, free-living

Introduction

The need for valid and reliable tools to accurately measure daily energy expenditure is an important public health research objective (23), particularly given the current epidemic of overweight and obesity. Numerous methods are available but each has limitations; methods with good validity tend to be too costly or complicated for widespread use; however, practical and feasible methods for large populations are limited by poor accuracy and/or reliability. The doubly labeled water (DLW) method is the “gold standard” for energy expenditure assessment under free-living conditions. The DLW method allows for assessment of total energy expenditure (TEE) over one to three weeks and has been shown to provide valid estimates of daily expenditure (27). The high cost and complicated analyses limits the use of DLW in most large epidemiological studies; however, the DLW technique provides a useful criterion measure for validating other instruments.

Accelerometers are the most practical and effective compromise between accuracy and feasibility for measurement of energy expenditure. They provide objective information about physical activity, are relatively inexpensive and are well tolerated by research participants. They have been widely used in research studies including use as a surveillance measure in the NHANES (24). Despite their wide acceptance in the research community, there are a number of limitations associated with the use of traditional accelerometry-based devices. There are currently many competing accelerometers and there is considerable confusion over the appropriate interpretation of accelerometer counts and the conversion of these counts into estimates of physical activity or energy expenditure (14, 25). Accelerometry-based activity monitors are also plagued by

challenges associated with detecting and addressing compliance (12). Non-wear time can often not be distinguished from periods of inactivity thereby necessitating detailed screening of data (2). While this detailed screening is common practice now, there is still considerable confusion about what constitutes a full day of monitoring or what counts as inactivity. Progress has been made to resolve these issues but it is likely that the fundamental challenges cannot be fully resolved without considerable enhancements in sensor and monitor technology.

Pattern-recognition monitors that integrate information from multiple sensors and have a notion of "context" have recently emerged as a possible advance for activity monitoring. The SenseWear Pro3 Armband (SWA; BodyMedia Inc., Pittsburgh, PA) is an example of such a pattern recognition device that addresses many of the limitations associated with single-axis accelerometers. The SWA monitor integrates information from a biaxial accelerometer and other physiological sensors (heat flux, temperature, and galvanic skin response sensors) to provide estimates of energy expenditure. The combination of sensors has shown to provide increased sensitivity for detecting the subtle changes in energy expenditure associated with complex lifestyle tasks and the increased energy expenditures associated with carrying loads, walking up grades or doing non-ambulatory activities (26). Previous studies have reported good validity of the SWA under laboratory conditions (4, 6, 7) but fewer studies have evaluated the device under free living conditions. One recent free-living study (21) found reasonable agreement between the SWA (Software version 4.02) and DLW for measurement of TEE in healthy adults, however significant differences in mean TEE between the two methods were found.

The proprietary algorithms for the SWA were recently modified, and the most current version of the software (Software V.6.1) has not been tested under free-living conditions. A smaller and thinner version of the SWA known as the SenseWear Mini was also recently developed by the same company. The Mini is based on the same technology but a three-axis accelerometer is used instead of the two-axis version in the SWA. The internal algorithms are slightly different between the monitors so validation of this monitor is also needed. The purpose of the present study was to evaluate the validity of the SWA and the Mini for measuring total daily energy expenditure under free-living conditions by direct comparison with doubly labeled water.

Methods

Participants

Thirty participants (15 male, 15 female) completed the study. The majority of the participants were Caucasian (73%) with 20% Hispanic and 7% Asian. Approximately 27% were overweight (BMI > 25) and 10% were characterized as obese (BMI > 30). Participants were between the ages of 24 and 60, did not have major disease or illness, did not use medications that would affect their body weight or metabolism, and were non-smokers. Participants were recruited through word of mouth. Approval from the Institutional Review Board of Iowa State University was obtained before beginning the study. Participants were aware of procedures and purposes of the study before they signed the informed consent document.

Instruments

The SenseWear Pro3 is a wireless multi-sensor activity monitor that integrates motion data from a two axis accelerometer along with several other physiological sensors

(heat flux, skin temperature and galvanic skin response). The monitor is worn on the upper arm over the tricep muscle and is lightweight (83-grams) and comfortable to wear. The data were processed using the latest proprietary algorithms available in the software (Software V.6.1, algorithm V2.2.3). The SenseWear Mini is a newer and smaller version of the SWA. The Mini operates in a similar manner but includes a triaxial accelerometer rather than a two axis accelerometer. Data were processed using similar algorithm architectures but with software specific for the Mini (Software V.7.0, algorithm V.2.2.4).

The software calculates the energy expenditure for each minute of data using complex pattern recognition algorithms, comprising of "activity classification" (context detection) and "energy expenditure estimation". A Naive Bays classifier is used to match the armband data to the activity class that best describes the current minute (the main classes are: walking, running, stationary bike, road bike, rest, resistance, and other activity). Each activity class has a linear regression model, mapping the sensor values and body parameters to energy expenditure. Kilocalories and metabolic equivalents (METs) are converted using the equation $\text{METs} = \text{kcal} / \text{hour} / \text{kg}$. The input to the Naive Bays classifier and the regression models include the data recorded in the armband and the standard deviation of the data over a number of minutes before and after the minute in question.

Data Collection Procedures

Anthropometric Data

Participants reported to the research center on the first day of the study (*day 1*) following a 10-h overnight fast (nothing to eat or drink except water). Anthropometric measurements were taken in light clothing and without shoes. Body weight was measured

to the nearest 0.1 kg with an electronic scale (Cardinal Detecto, Webb City, MO) and height was measured to the nearest 0.1 cm with a wall-mounted stadiometer (Ayrton, Prior Lake, MN) Body mass index (BMI) was calculated as weight (kg)/height² (m²).

Activity Monitor Data

The monitors were initialized using the participant's personal information (age, gender, height, weight) and adjusted to fit on the participant's arm. The SWA monitor was placed on the right arm and the Mini monitor was placed on the left arm, according to manufacturer recommendations. Participants were instructed to wear both monitors simultaneously from day 1 until day 14 (including while sleeping) but were allowed to remove the monitors briefly for showers or water activities. Participants were asked to keep a diary of non-wearing periods. Careful attention was given to processing the individual files to ensure that any gaps in the data did not influence the results. The brief gaps in the data attributable to showering were manually filled with a corresponding MET equivalent for "self-care activities" (2.0 METs) based on the Compendium of Physical Activities (1). The software detected other shorter gaps in the data which occur if the strap is loosened or jostled during daily activities. The monitor beeps when this occurs, reminding the participant to tighten/adjust the strap, but small gaps still occur. These gaps (ranging in duration from 1-8 minutes) were filled with the average energy expenditure of the 10 minutes before the defined gap. Planned gaps in the data occurred on Day 7 for download and battery change/charge. This time was filled using MET estimates of light/resting activity since participants completed a sedentary activity during this time.

Several other longer periods of non-wear time (> 30 minutes) were detected in the downloaded files. These gaps were filled with 2.0 MET values, except in cases where participants reported doing water activities, for which corresponding MET values were applied using the Compendium of Physical Activities (1). One participant removed the monitor while sleeping (547 minutes) so this gap was filled with estimates from sleeping on other nights. A large gap of over 1529 minutes (greater than 1 day) was also detected for one participant and this occurred due to a battery malfunction. For this case, we used the corresponding values from the same day on the previous week to fill the data. The overall compliance with the monitoring protocol was excellent. The measured wear time (including all of the above potential gaps) averaged 1401 minutes per day ($97.1 \pm 3.6\%$). The average non-wear time (about 39 minutes) was biased to some extent by the few participants that had larger gaps in the data. When the large gaps for the sleep and battery malfunction were excluded, the average non-wear time was 33 minutes per day.

Doubly Labeled Water Data

The DLW technique provided criterion measures of TEE. Following the collection of two baseline (*day 1*) urine samples, participants were dosed with 1.5 ml/kg body wt of a mixture of 10% enriched H_2^{18}O and 99% enriched $^2\text{H}_2\text{O}$ (Cambridge Isotopes, Cambridge, MA). The dose was followed by a 100-ml tap water rinse to ensure complete delivery of the labeled water. The first two urine samples after dosing (~1.5 and 3 h postdose) were discarded followed by two urine samples collected at 4.5 h and 6.0 h after dosing. On the mornings of day 7 and day 14, participants were instructed to discard their first urine void and collect the second void of the day. Samples were collected in airtight containers and were brought to the research center in cooler packs. Abundance of

^{18}O was measured in duplicate on a Finnigan MAT 252 dual inlet gas isotope ratio mass spectrometer (IRMS), and $^2\text{H}_2$ abundance was measured in duplicate on the same IRMS using a Finnigan H/D equilibration device. The ^2H and ^{18}O isotope elimination rates (k_{D} and k_{O}) were calculated using linear regression following a log transformation. Total body water (N) was determined at time zero, obtained from the regression line of the H_2^{18}O isotope. The rate of CO_2 production was calculated using the equations of Schoeller et al. (19) and later modified (17) as follows:

$$r\text{CO}_2 \text{ (moles/d)} = (N/2.078) (1.007k_{\text{O}} - 1.041k_{\text{D}}) - 0.0246r_{\text{GF}};$$

where $r\text{CO}_2$ is the rate of carbon dioxide production; N is total body water calculated from $N_{\text{O}}/1.007$ where N_{O} is the ^{18}O dilution space; k_{O} and k_{D} represent the fractional elimination rates of ^{18}O and $^2\text{H}_2$, respectively; and r_{GF} is the rate of fractionated gaseous evaporative water loss, which is estimated to be $1.05 \cdot N (1.007k_{\text{O}} - 1.041k_{\text{D}})$. Total energy expenditure (TEE) was calculated as follows: $\text{TEE (kcal/d)} = 22.4 r\text{CO}_2 (3.9/\text{RQ} + 1.10)$. This formula assumes a respiratory quotient (RQ) of 0.86 which is typical for a healthy, rather low fat diet. The corresponding energy equivalent of CO_2 (Eeq_{CO_2}) was 5.637 kcal/L CO_2 . The DLW was processed and analyzed at the Pennington Biomedical Research Center. The intra-assay variability for DLW assessments is less than 2%.

To assess the ability of the monitors to measure energy expenditure associated with physical activity, we calculated daily physical activity energy expenditure (PAEE) using the following equation (18): $\text{PAEE} = \text{TEE} - (\text{resting metabolic rate} + 0.1 \cdot \text{TEE})$. This approach uses the standard assumption that the thermic effect of food is approximately 10% of TEE (22). The resting metabolic rate (RMR) in this equation was estimated using standard WHO equations (20) which are based on weight, age and

gender. The PAEE provides an indicator of total activity level but we also calculated standardized estimates of usual physical activity level (PAL), calculated as TEE from DLW divided by estimated RMR.

Statistical Analyses

This study evaluated the agreement between estimates of TEE from the SWA and the Mini compared with criterion estimates from the DLW. A secondary analysis included the comparison of PAEE estimates between the monitors and DLW. Primary statistical analyses were performed using JMP software v.7.0 (SAS Institute, Cary, NC). Paired *t* tests were used to determine differences between the mean values obtained with the monitors and DLW. Simple linear regression analyses were conducted between armband and DLW energy expenditure estimates to evaluate the associations between measures. Analysis of covariance (ANCOVA) was used to test for the effects of gender on estimates of energy expenditure. No significant effects of gender were found; therefore, males and females were combined for all analyses. To evaluate the extent of agreement between measures of energy expenditure, intra-class correlations (ICC, one-factor random effect) were computed to determine agreement between measures (correlations closer to 1.0 indicate greater agreement) (3).

Graphical procedures were used to examine agreement across the range of TEE values (graphed using MATLAB v7.1, MathWorks, Natick, MA). The DLW estimate of TEE was plotted on the x-axis (rather than the mean of the DLW and monitor scores) because the DLW is a criterion measure. The differences between the DLW and monitor estimates were plotted on the y-axis to demonstrate the individual variability in responses. Confidence intervals defining the limits of agreement were established as 1.96

SD from the mean difference. To evaluate the presence of systematic bias, residuals of armband estimates of EE were plotted against the reference method DLW.

Results

Physical characteristics for the 30 participants are presented in Table 1 along with descriptive PAL measures from the DLW measurements and RMR estimates. Table 2 contains the TEE and PAEE estimates from the monitors and differences from the DLW values. The SenseWear Pro3 tended to underestimate TEE compared to the DLW method; however, the difference in TEE between the two methods was non-significant ($p = 0.07$). The SWA was lower by an average of 112 kcal/d compared to DLW, representing an average of 4% underestimation in TEE. There was no significant difference in TEE estimates ($p=0.69$) between the SenseWear Mini monitor and DLW. The Mini was lower by an average of 22 kcal/d compared to DLW, an average underestimation of <0.1% in TEE. Estimates of TEE from the SWA and the Mini were not significantly different from each other ($p=0.5$). While the difference from DLW was smaller with the Mini, the error rates were similar (SWA: $8.1 \pm 6.8\%$, Mini: $8.3 \pm 6.5\%$) when expressed as absolute percent error (computed using the absolute value of the differences). This suggests that the two monitors had similar absolute error compared to the DLW method. The regression analyses showed significant agreement between the SWA and DLW measurements of TEE ($R^2 = 0.68$, $p<0.001$) and the Mini and DLW ($R^2 = 0.71$, $p<0.001$) (Figure 1).

Intraclass correlations were used to examine individual agreement in TEE values between the armband monitors and DLW. The ICC for the SWA and DLW was 0.80 (95% CI: 0.89, 0.70), indicating that 80% of the variance in the measurements was

explained by differences between individuals, whereas 20% was due to variation in the two methods. The ICC for the Mini and DLW was of 0.85 (95% CI: 0.92, 0.76), suggesting that 85% of the variation in TEE estimates was due to differences between individuals and 15% was due to variation between the Mini and DLW methods. For both monitors, the ICCs exceeded the generally accepted threshold for good agreement of 0.75 (3).

The graphical plots in Figure 2 provide a more detailed view of the differences in measurement agreement between the monitors and the DLW method. The plots examine comparisons between the armbands and DLW by plotting differences in total daily EE between DLW and the armbands versus mean daily EE determined from DLW. Limits of agreement from the plots (mean \pm 1.96 SD) were slightly smaller with the Mini monitor (-630 to 585 kcal/d) than with the SWA monitor (-749 to 525 kcal/d). Results of regressing the difference in TEE between each monitor and DLW (monitor TEE – DLW TEE) against DLW measures of TEE indicate that both monitors overestimated TEE at low levels of daily energy expenditure and significantly underestimated TEE at higher levels (SWA: $R^2 = 0.30$, $p=0.002$, Mini: $R^2 = 0.19$, $p=0.02$) (Figure 3). Overall, the SWA underestimated TEE in 19 of the 30 participants (63%) and the Mini underestimated TEE in 17 of the 30 participants (57%).

As a secondary analysis, we also examined the agreement between estimates of PAEE from the monitors compared to DLW. Both monitors significantly underestimated PAEE compared to DLW estimates (SWA: $p=0.02$, Mini: $p=0.03$, Table 2). The ICCs for both monitors with DLW PAEE were 0.63 (95% CI: 0.77, 0.47), indicating that 63% of the variance in the measurements was explained by differences between individuals

and 37% was due to variation between the methods. Regression analyses revealed modest agreement between SWA PAEE and DLW PAEE ($R^2 = 0.51$, $p < 0.001$) and similar agreement between Mini PAEE and DLW PAEE ($R^2 = 0.48$, $p < 0.001$). Regressing the residual values of activity energy expenditure (monitor PAEE – DLW PAEE) on DLW PAEE revealed that both monitors significantly under-estimated energy expenditure at higher levels of PAEE (SWA: $R^2 = 0.56$, $p < 0.001$, Mini PAEE: $R^2 = 0.49$, $P < 0.001$). Absolute error of PAEE estimates was 26% for the SWA monitor and 28% for the Mini monitor, compared with DLW measures.

Due to the rather large ranges of age, BMI, and physical activity level across the participants, we conducted additional correlation analyses to determine if the residuals from TEE and PAEE were related to these variables. We found no significant relationships between the TEE residuals and age or BMI; however, more negative residuals were associated with a higher PAL for both monitors (SWA, $R^2 = 0.19$, $p = 0.02$; Mini, $R^2 = 0.18$, $p = 0.02$), again suggesting that the monitors underestimate TEE at higher levels of energy expenditure. Residuals in PAEE were not associated with age, BMI, or PAL.

Discussion

We evaluated the accuracy of two armband monitors for measuring TEE and PAEE in healthy adults under free-living conditions compared with the gold standard doubly labeled water. The development of accurate, reliable, and affordable tools to measure daily energy expenditure in free-living conditions is an important priority for public health researchers. The results of the present study support the use of these monitors for estimated daily energy expenditure but the recently developed Sensewear

Mini showed slightly better performance over the SenseWear Pro3, possibly due to the inclusion of a three-axis (versus two-axis) accelerometer. For both monitors, examination of the graphical plots of the data and analyses of the residual vs. DLW TEE values revealed greater underestimation of energy expenditure at higher TEE.

The results from the present study are generally consistent with findings from a previous study (21) that evaluated an earlier version of the SWA (software v.4.2). St-Onge et al. (21) reported that the SWA underestimated TEE by 117 kcal/day ($p < 0.01$) compared to DLW over ten days, whereas we found an underestimation of 112 kcal/d. They reported similar values to ours for agreement between the SWA and DLW measures ($ICC = 0.81$) and association between the methods ($R^2 = 0.74$). Also consistent with our findings, was the observation that the monitor under-estimated TEE at higher levels of expenditure ($R^2 = 0.33$). Our study is the first to report data on the Mini monitor, which suggests some improvement in estimating TEE over the older SWA versions.

A number of other studies have investigated the accuracy of activity monitors for measuring TEE compared to DLW, and these were highlighted in a recent review (16). The monitor most frequently studied is the Actigraph (Actigraph, Pensacola, FL), formerly known as the CSA (Computer Science Applications) and the MTI (Manufacturing Technology Inc., Fort Walton Beach, FL). Validation studies of this monitor against DLW have involved primarily women, with additional studies in children and adolescents (16). Lof et al. (11) compared TEE measured by the CSA with DLW in 34 women and found a non-significant difference of 88 kcal/d; however, the limits of agreement were large (± 700 kcal/d). Another study by this group (10) again compared

the CSA to DLW in 24 healthy women and found a mean difference in TEE of 6 ± 325 kcal/d, with no systematic bias present. Masse et al. (13) compared the CSA monitor to DLW in a group of 136 African American and Hispanic women. The monitor was worn during the last 7 days of the 14-day DLW period. Modest agreement was found between the monitor and DLW estimates of TEE, ranging from $R^2 = 0.39$ to 0.44 , depending on whether controlling for body mass or fat-free mass. Leenders et al. (9) compared TEE measured by the CSA and Tritrac monitors to DLW and found modest associations (R^2 from 0.17 to 0.45); however, concordance between the monitors and DLW ranged from 0.04 to 0.50 , considered slight to moderate (8). Both monitors provided reasonable estimates of TEE with some underestimation (-2 to -23%), however standard deviations were large.

A study by Hustvedt et al. (5) evaluated the accuracy of the ActiReg (a three-dimensional accelerometer) alone and in combination with a heart rate monitor. Mean TEE measured by the ActiReg was not different from DLW ($p=0.45$), with a mean difference of 98 kcal/d and limits of agreement of -397 to 765 kcal/d. Bland-Altman plots showed that the ActiReg underestimated TEE at higher levels of energy expenditure, which was reduced by using the heart rate function. Plasqui et al. (15) evaluated another three-dimensional monitor called the Tracmor. They reported that participant characteristics such as age, body mass, and height explained 64% of the variation in DLW TEE but adding Tracmor activity counts to the model increased the explained variation by 19% (total $R^2 = 0.83$). In our study, we found that age, body mass and height explained 40% of the variation in DLW TEE. Adding SWA TEE to the model increased the explained variation by 29% (total $R^2 = 0.69$, $p<0.001$) indicating that the

movement and other physiological sensors provide useful information to improve estimates of TEE. Similar results were obtained when the analyses were repeated with the Mini with the explained variance increasing by 32% ($R^2 = 0.72$, $p < 0.001$). Adding both monitors to the model did not explain further variance, indicating that the monitors function independently and there is no additive benefit of using multiple monitors.

We found poorer performance of the monitors for measuring PAEE. Both monitors significantly underestimated PAEE (particularly at higher levels of activity expenditure) and may have contributed to the overall underestimation of TEE. St-Onge et al. (21) reported an average underestimation in PAEE of 225 kcal/d, while we observed an underestimation of 123 and 119 kcal/d (SWA and Mini, respectively), suggesting some improvement. We also found better agreement between the monitors and DLW (ICC = 0.63) than the St-Onge study (ICC = 0.46) despite similar associations between the methods ($R^2 = 0.48$ and 0.51 vs. St-Onge $R^2 = 0.49$). It is important to note, however, that caution should be exercised when interpreting the results of our PAEE data, as RMR was not measured with an indirect calorimeter but instead was estimated using WHO equations (20).

The results of the present study show that the SenseWear Pro3 and the Sensewear Mini perform similar to or better than other available monitors. Both monitors underestimated energy expenditure, particularly at higher levels of expenditure, and this underestimation continues to be a problem among many activity monitors that are currently available. The tendency for underestimation can be attributed to the inherent challenges of capturing low intensity activities of daily living, which contribute to TEE but are difficult to detect with accelerometer technology. An advantage of the multi-sensor

armband technology is the inclusion of thermal and perspiration related sensors as well as accelerometers. The heat sensors provide a way to detect the subtle increases in energy expenditure associated with low intensity activities. Previous research with the armband has indicated that the SWA provides more accurate estimates of low intensity activities than the Actigraph (26). This may have contributed to the improved performance relative to previous DLW studies with the Actigraph. Another possible reason for the improved result is the better detection (and correction) of non-wear time. This capability removes the guess-work that is often needed to address gaps in the data that occur with other monitors, allowing for more confidence in the results (i.e., under- or over-estimation by the monitors is due to capability of the monitor and not to error introduced by using assumed energy expenditure data).

Our study is not without limitations. Most volunteers who participated were quite lean and active, likely more so than an average population. There is also the possibility of “reactivity” when wearing the monitors; i.e., subjects may have increased their daily activity over their usual patterns due to wearing the monitor. However, our subjects were instructed to maintain their usual daily routines of work, activity, etc. while wearing the monitors. Also, limitations of the monitors themselves include the proprietary nature of the algorithms, which do not allow for independent investigators to work with the algorithms, and the cost of the monitor, which is less than some available but is likely not a negligible expense.

In summary, the SenseWearPro3 and the SenseWear Mini armbands show promise for accurately measuring daily energy expenditure under free-living conditions. An advantage of these monitors is that they provide direct estimates of wear time and

avoid challenges associated with evaluating compliance. The monitors also provide direct estimates of energy expenditure and avoid the confusion in the literature caused by the availability of different calibration equations for different populations. However, more work is needed to improve the ability of these monitors to accurately measure energy expenditure at higher levels of expenditure.

Disclosures

This research was funded by a grant from Bodymedia Inc. awarded to Dr. Greg Welk.

Acknowledgments

The results of the present study do not constitute endorsement by ACSM.

References

- 1- Ainsworth BE, Haskell WL, Whitt MC et al. Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc.* 2000; 32 (9 Suppl):S498-504.
- 2- Catellier DJ, Hannan PJ, Murray DM et al. Imputation of missing data when measuring physical activity by accelerometry. *Med Sci Sports Exerc.* 2005; 37 (11 Suppl):S555-62.
- 3- Fleiss J. Reliability of measurements. In, translator and editor The design and analysis of clinical experiments. New York: John Wiley & Sons; 1986; p. 2 - 31.
- 4- 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.
- 5- Hustvedt BE, Christophersen A, Johnsen LR et al. Description and validation of the ActiReg: a novel instrument to measure physical activity and energy expenditure. *Br J Nutr.* 2004; 92 (6):1001-8.
- 6- Jakicic JM, Marcus M, Gallagher KI et al. Evaluation of the SenseWear Pro Armband to assess energy expenditure during exercise. *Med Sci Sports Exerc.* 2004; 36 (5):897-904.
- 7- King GA, Torres N, Potter C et al. Comparison of activity monitors to estimate energy cost of treadmill exercise. *Med Sci Sports Exerc.* 2004; 36 (7):1244-51.
- 8- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977; 33 (1):159-74.
- 9- Leenders NY, Sherman WM, Nagaraja HN. Energy expenditure estimated by accelerometry and doubly labeled water: do they agree? *Med Sci Sports Exerc.* 2006; 38 (12):2165-72.
- 10- Lof M, Hannestad U, Forsum E. Assessing physical activity of women of childbearing age. Ongoing work to develop and evaluate simple methods. *Food Nutr Bull.* 2002; 23 (3 Suppl):30-3.
- 11- Lof M, Hannestad U, Forsum E. Comparison of commonly used procedures, including the doubly-labelled water technique, in the estimation of total energy expenditure of women with special reference to the significance of body fatness. *Br J Nutr.* 2003; 90 (5):961-8.
- 12- Masse LC, Fuemmeler BF, Anderson CB et al. Accelerometer data reduction: a comparison of four reduction algorithms on select outcome variables. *Med Sci Sports Exerc.* 2005; 37 (11 Suppl):S544-54.
- 13- Masse LC, Fulton JE, Watson KL et al. Influence of body composition on physical activity validation studies using doubly labeled water. *J Appl Physiol.* 2004; 96 (4):1357-64.
- 14- Matthew CE. Calibration of accelerometer output for adults. *Med Sci Sports Exerc.* 2005; 37 (11 Suppl):S512-22.
- 15- Plasqui G, Joosen AM, Kester AD et al. Measuring free-living energy expenditure and physical activity with triaxial accelerometry. *Obes Res.* 2005; 13 (8):1363-9.
- 16- Plasqui G, Westerterp KR. Physical activity assessment with accelerometers: an evaluation against doubly labeled water. *Obesity (Silver Spring).* 2007; 15 (10):2371-9.

- 17- Racette SB, Schoeller DA, Luke AH et al. Relative dilution spaces of 2H- and 18O-labeled water in humans. *Am J Physiol.* 1994; 267 (4 Pt 1):E585-90.
- 18- Rising R, Harper IT, Fontvielle AM et al. Determinants of total daily energy expenditure: variability in physical activity. *Am J Clin Nutr.* 1994; 59 (4):800-4.
- 19- Schoeller DA. Measurement of energy expenditure in free-living humans by using doubly labeled water. *J Nutr.* 1988; 118 (11):1278-89.
- 20- Schofield WN. Predicting basal metabolic rate, new standards and review of previous work. *Hum Nutr Clin Nutr.* 1985; 39 Suppl 1:5-41.
- 21- St-Onge M, Mignault D, Allison DB et al. Evaluation of a portable device to measure daily energy expenditure in free-living adults. *Am J Clin Nutr.* 2007; 85 (3):742-9.
- 22- Tataranni PA, Larson DE, Snitker S et al. Thermic effect of food in humans: methods and results from use of a respiratory chamber. *Am J Clin Nutr.* 1995; 61 (5):1013-9.
- 23- Troiano RP. A timely meeting: objective measurement of physical activity. *Med Sci Sports Exerc.* 2005; 37 (11 Suppl):S487-9.
- 24- Troiano RP, Berrigan D, Dodd KW et al. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc.* 2008; 40 (1):181-8.
- 25- Welk GJ. Principles of design and analyses for the calibration of accelerometry-based activity monitors. *Med Sci Sports Exerc.* 2005; 37 (11 Suppl):S501-11.
- 26- Welk GJ, McClain JJ, Eisenmann JC et al. Field validation of the MTI Actigraph and BodyMedia armband monitor using the IDEEA monitor. *Obesity (Silver Spring).* 2007; 15 (4):918-28.
- 27- Westerterp KR, Brouns F, Saris WH et al. Comparison of doubly labeled water with respirometry at low- and high-activity levels. *J Appl Physiol.* 1988; 65 (1):53-6.

Tables

Table 1. Participant characteristics [mean \pm SD and range (min – max)]

	Men (n=15)	Women (n=15)	All (n=30)
Age (y)	36.3 \pm 9.7 (25 – 51)	40.1 \pm 11.5 (24 – 60)	38.2 \pm 10.6
Height (cm)	176 \pm 6 (166 – 188)	167 \pm 4 (162 – 175)	171 \pm 7
Weight (kg)	78.5 \pm 11.3 (66.1 – 107.7)	63.6 \pm 11.6 (50.4 – 94.0)	71.2 \pm 13.7
BMI (kg/m²)	25.4 \pm 3.0 (20.8 – 30.6)	22.7 \pm 3.3 (17.8 – 31.5)	24.0 \pm 3.4
Physical Activity Level	1.7 \pm 0.3 (1.2 – 2.4)	1.8 \pm 0.2 (1.5 – 2.4)	1.73 \pm 0.3

Table 2. Comparisons of total daily energy expenditure [mean \pm SD]

	Mean \pm SD	Mean difference (variable – DLW)	Range
Total Daily EE (kcal/day)			
Doubly Labeled Water	2,774 \pm 576		2,009 – 4,144
SenseWear Pro3	2,662 \pm 483	-112 \pm 325*	1,798 – 3,664
SenseWear Mini	2,752 \pm 523	-22 \pm 310	1,764 – 3,868
Physical Activity EE (kcal/day)			
Doubly Labeled Water	983 \pm 486		503 – 1515
SenseWear Pro3	769 \pm 265	-123 \pm 278 †	242 – 1,379
SenseWear Mini	773 \pm 283	-119 \pm 286 †	254 – 1,575

EE, energy expenditure

* p=0.07

† Significantly different from doubly labeled water, p<0.05

Figures**Figure 1.**

Regression analysis with Pearson's Correlation Coefficients between the armbands and doubly labeled water methods for measuring daily energy expenditure ($n = 30$, $p < 0.001$ for both).

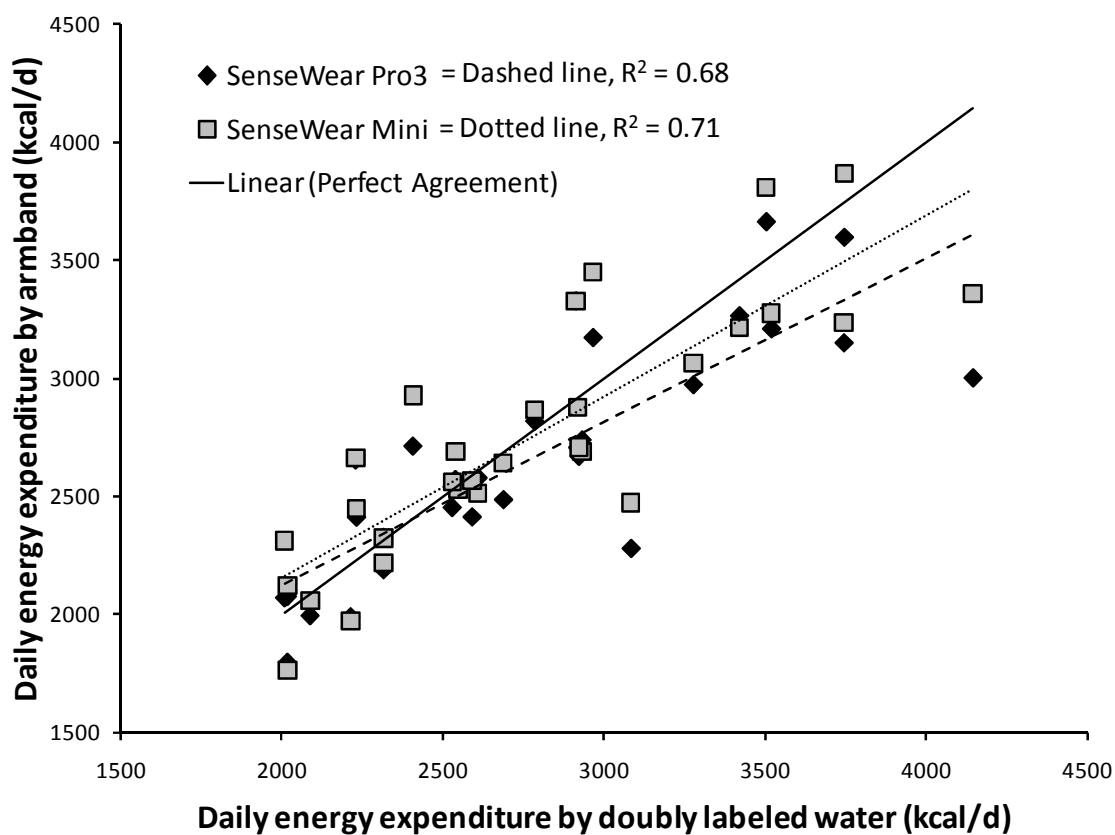
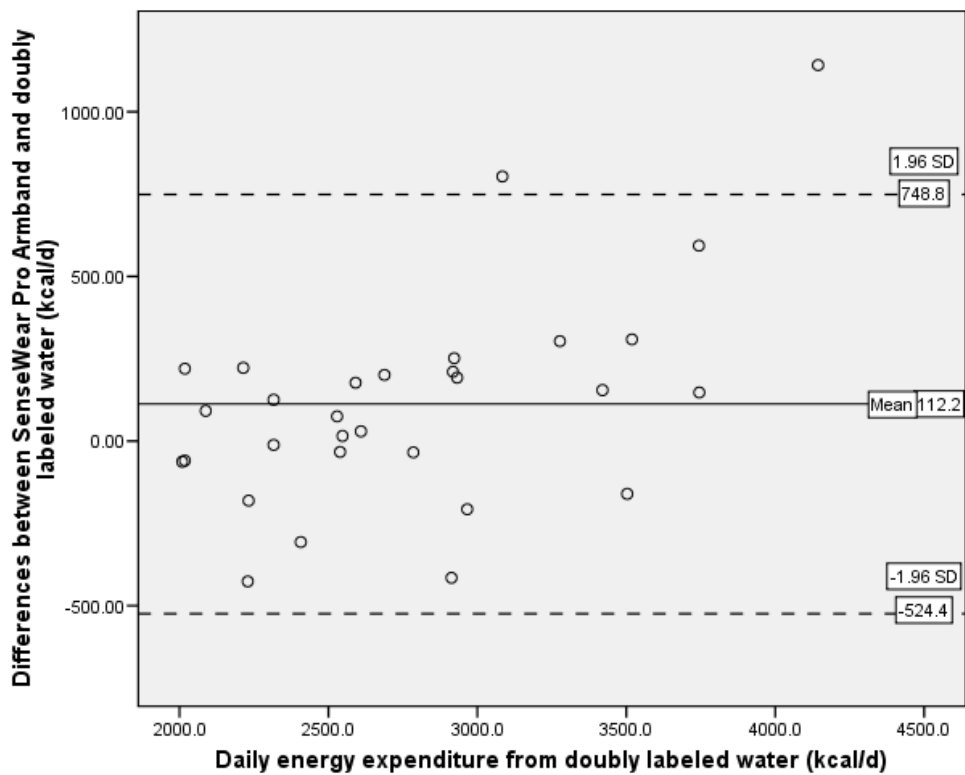


Figure 2.

Top panel. Graphical plot between the SenseWear Mini Armband and doubly labeled water methods for measuring daily energy expenditure (n = 30). Estimates were averaged across 14 days to provide one estimate of energy expenditure.

Bottom panel. Graphical plot between the SenseWear Pro3 Armband and doubly labeled water methods for measuring daily energy expenditure (n = 30). Estimates were averaged across 14 days to provide one estimate of energy expenditure.



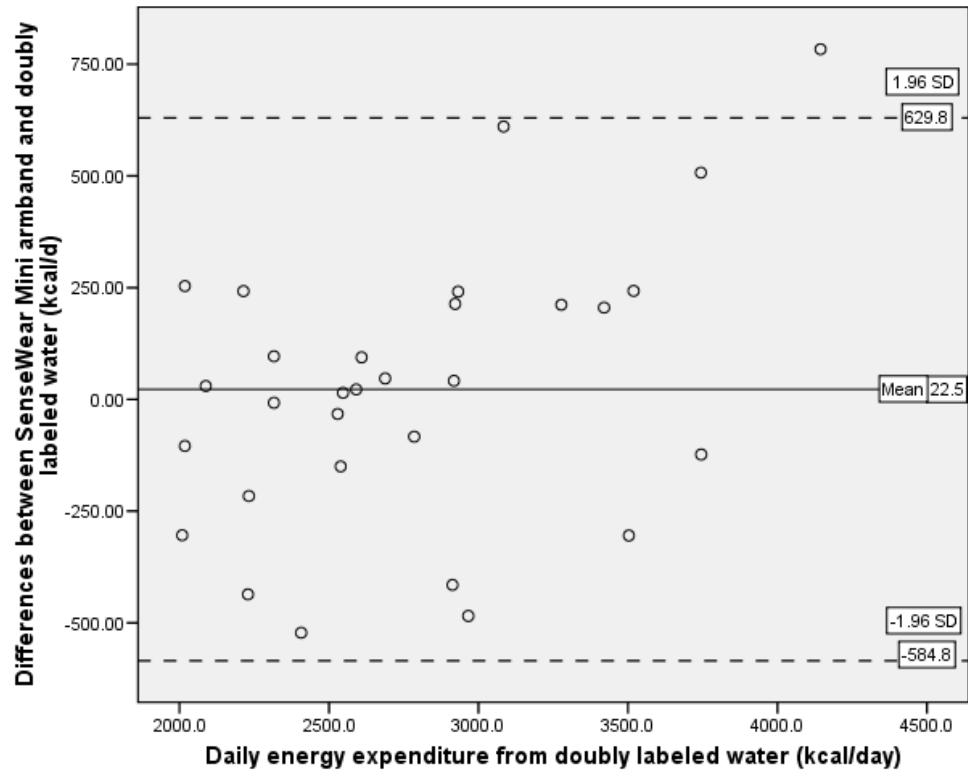
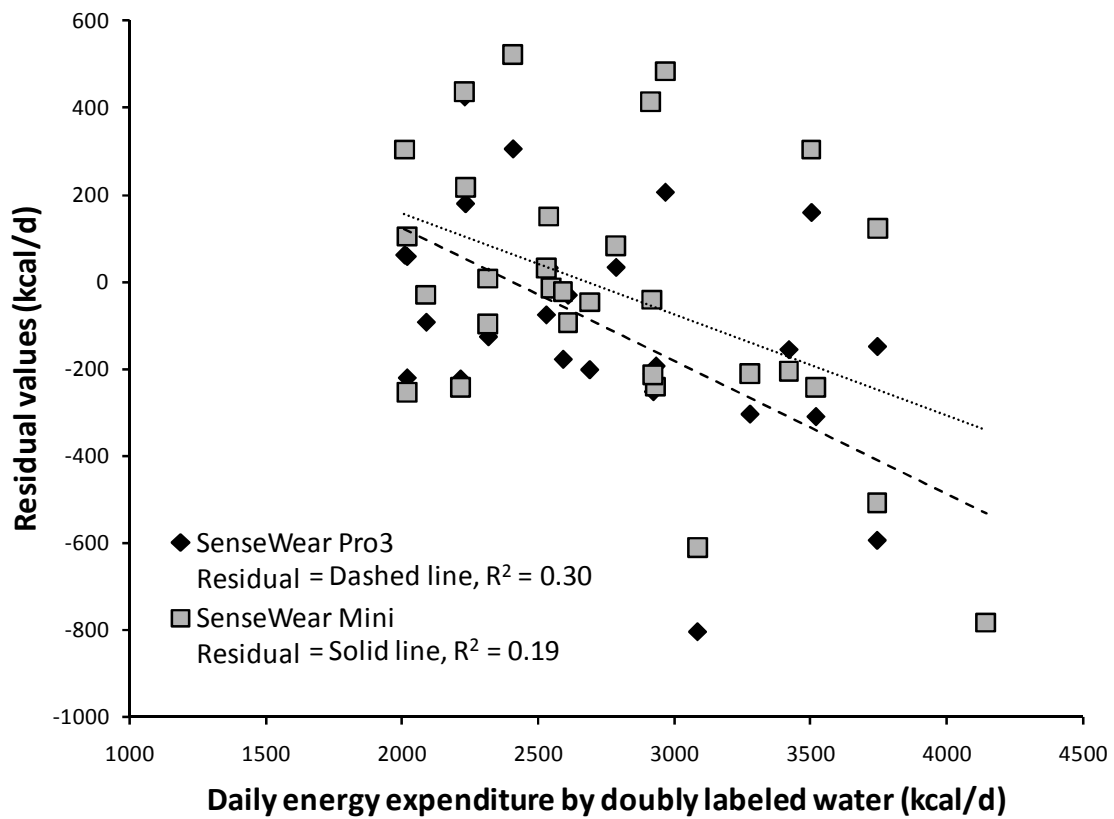


Figure 3.

Residual values for daily energy expenditure (EE) plotted against the reference (doubly labeled water, DLW) method for measuring daily EE ($n = 30$). R^2 for the SWA regression is 0.30, $p=0.002$, and R^2 for the Mini is 0.19, $p=0.02$, indicating that both monitors significantly under-estimate TEE at a higher DLW TEE.

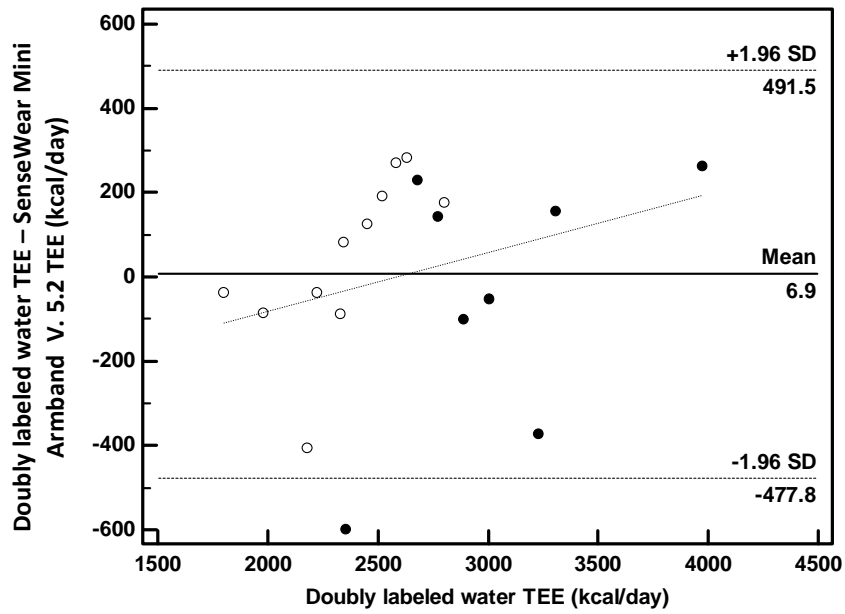


APPENDIX B

Figure 1- New SenseWear Armband algorithms (5.2)

Gender reference: ● Males, ○ Females

a) SenseWear Mini Armband 5.2 (Mini 5.2)

b) SenseWear Pro₃Armband 5.2 (SWA 5.2)

