2018

Estimation of physical activity using accelerometry in adult populations: Using the Sensewear Armband and the ACT24 as comparison tools for the estimation of energy expenditure and physical activity intensity by the Sojourn method

Matthew Stewart
Iowa State University

Estimation of physical activity using accelerometry in adult populations:

Using the Sensewear Armband and the ACT24 as comparison tools for the estimation of energy expenditure and physical activity intensity by the Sojourn method

by

Matthew Stewart

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Kinesiology

Program of Study Committee:
Gregory Welk, Major Professor
Laura Ellingson
Christina Campbell

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this thesis. The Graduate College will ensure this thesis is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2018

TABLE OF CONTENTS

Page

LIST OF TABLES iii

LIST OF FIGURES iv

ABSTRACT v

CHAPTER 1. INTRODUCTION 1

CHAPTER 2. REVIEW OF LITERATURE 4
- Overview of PA and SB Assessment Methods 6
  - Report-Based Methods 6
  - Monitor-Based Methods. 10
- Accelerometers: A Closer Look 11
  - Actigraph. 12
  - Sensewear Armband. 14
- Overview of Sojourn Method 17

CHAPTER 3. METHODS 22
- Overview of Study and Design 22
- Instruments 22
  - Sensewear Core Armband (SWA) 22
  - Actigraph wGT3X-BT 23
  - ACT24 23
- Data Collection 23
  - First Appointment 24
  - Second Appointment 24
- Data Processing 25
- Data Analysis 27
  - Descriptive Analyses 27
  - Method Agreement Analyses 27
  - Error Visualization 28

CHAPTER 4. RESULTS 29

CHAPTER 5. DISCUSSION 35
- Summary, Limitations, and Conclusions 41

TABLES AND FIGURES 46

REFERENCES 56

LIST OF TABLES

Table 1. Baseline Demographic and General Health Characteristics 46

Table 2. Average Activity Level by Method and Sex 48

Table 3. Time Spent at Various Intensities: Sojourns and SWA 50

Table 4. Time Spent at Various Intensities: SWA and ACT24 50

Table 5. Time Spent at Various Intensities: Sojourns and ACT24 50

Table 6. EE Correlation 50

Table 7. MPE and MAPE Analyses 51

LIST OF FIGURES

Figure 1. Average Activity Level by Intensity Classification 52

Figure 2. Bland-Altman Plots: Energy Expenditure Estimation..53

Figure 3. Bland-Altman Plots: Sojourns and SWA 54

Figure 4. Bland-Altman Plots: Sojourns and ACT24 55

ABSTRACT

Accurate assessments of physical activity are essential for advancing many lines of physical activity research. Numerous physical activity assessment techniques have been developed, but continual refinement and evaluation of these techniques is important to further improve accuracy and precision. A machine-learning technique known as the Sojourn method has demonstrated promise for improving the accuracy of accelerometer-based physical activity monitors, such as the widely used Actigraph. However, fewer studies have validated this method under free-living conditions. **Purpose:** The purpose of this study was to examine the performance of the Sojourn method for estimating energy expenditure and time spent at various intensities of activity relative to estimates from an established monitor-based measure (Sensewear Armband) and an established report-based measure (ACT24). A secondary purpose is to examine the context-related factors that are captured with the Sojourn method by comparing activity patterns with parallel data from the Sensewear Armband (SWA) and ACT24. **Methods:** The study used data obtained through a large (ongoing) field-based evaluation of activity monitors conducted in the Physical Activity and Health Promotion Lab at Iowa State University. The data for the present study involved temporally matched data from a sample of 85 adults with complete data on these three measures (Sojourns, SWA, and ACT24). The study involved two laboratory assessment days split by a 24-hour period during which activity monitors were worn. The first meeting consisted of participants completing a demographic survey and anthropometric measurements. Participants were instructed to wear the Actigraph and the Sensewear monitor (along with 5 other monitors) for a full 24-hour period (midnight to midnight) under free-living conditions. On the day following monitor wear, participants returned to the lab to complete the ACT24. Correlations, mean percent error (MPE), mean absolute percent error (MAPE), and

Bland-Altman plots were used to assess method agreement. **Results:** Correlational analyses revealed moderate-strength relationships between the Sojourns and SWA (r = 0.65) and strong associations between the Sojourns and ACT24 (r = 0.91) for capturing total daily energy expenditure. Additionally, correlational analyses revealed moderate-strength relationships between the Sojourns and each of the other methods for time spent in sedentary, vigorous intensity activity, and MVPA. Error analyses revealed modest amounts of error between the Sojourns and SWA (MPE: 3.5%, MAPE: 16.1%) as well as between the Sojourns and ACT24 (MPE: 6.6%, MAPE: 9.5%) for capturing total daily energy expenditure. Error between methods was lowest for sedentary time (with all values below 24%) but classification accuracy was higher for light and moderate intensity activity. Bland-Altman analysis revealed some bias between methods for all indicators. **Conclusions:** The Sojourn method has promise as a standardized method for estimating energy expenditure and time spent in physical activity. However, additional refinements are warranted to further improve the utility for field-based research applications.

CHAPTER 1. INTRODUCTION

A key priority in contemporary physical activity research is to develop methods that can accurately assess physical activity behaviors under free-living conditions in the population. These estimates of physical activity and sedentary behavior are needed for a variety of applications in behavioral epidemiology. Surveillance research requires accurate assessment options to identify population-wide trends and disparities between characteristics such as age, sex, and ethnicity, among others. Such assessments are also needed to identify correlates of such behavior, which can then be built into interventions targeting physical activity or sedentary behavior and their health-related outcomes.

The most common tools for assessing physical activity behavior are report-based methods and monitor-based methods. Report-based measures include various self-report questionnaires, activity logs, and diaries. These measures are often easily administered, cheap, and have a relatively low burden on the participant. However, they can also be subject to questionable validity and accuracy (Prince et al., 2008). A popular alternative to report-based measures are various accelerometry-based measures that provide more objective estimates of physical activity behavior. These monitor-based methods do not rely on self-report, and thus are not subject to the same errors associated with recall or bias (Prince et al., 2008). However, a limitation is that objective measurement tools do not provide context on physical activity behavior. Another challenge of wearable monitors is that it is difficult to convert the data into meaningful and useable indicators, such as time spent at different intensities of activity or energy expenditure. To improve utility, researchers have developed calibration methods that convert the raw data into such outcomes. Several popular methods applied to hip-worn Actigraph data have been developed by Freedson, Pober, and Janz

(2005) and Crouter, Clowers, and Bassett (2006), but a number of others have also been developed for the Actigraph (Evenson, Catellier, Gill, Ondrak, & McMurray, 2008).

Despite the promise of obtaining robust data from wearable monitors, the wide variety of calibration methods available has hindered the interpretation and comparability of activity data across studies. There has not been a consensus on what calibration method should be used, forcing researchers to select from several options and leading to what some refer to as the cut-point conundrum (Trost, Loprinzi, Moore, & Pfeiffer, 2011).

A promising new calibration method, known as the Sojourn method, was recently developed by Lyden, Kozey-Keadle, Staudenmeyer, and Freedson (2014). This hybrid machine learning technique combines artificial neural networks with decision tree analysis. Used for free-living activities, this method uses counts per second from the signals of three different axes from the hip-mounted Actigraph accelerometer. It consists of three steps: identifying the number of bouts, determining whether these bouts are activity or inactivity, and assigning MET values to both inactivity and activity. This process is fully described later. An estimation of time spent at different intensities of activity as well as energy expenditure can be derived from this process. The method may also have certain advantages over other methods. For example, previous approaches typically segment the signal into windows of fixed length, called epochs (Gabriel et al., 2010). The artificial time boundaries of this approach has limitations when applied to activities that are unplanned and of different length (Lyden, Kozey-Keadle, Staudenmayer, & Freedson, 2014; Crouter, Kuffel, Haas, Frongillo, & Bassett, 2010). The Sojourn method avoids this issue by using nonfixed, activity-defined windows to segment the signals.

The original paper on the Sojourn method supported the approach (Lyden, Kozey-Keadle, Staudenmayer, & Freedson, 2014), but few studies have compared the Sojourn method against other available methods. Recently, Ellingson et al. (2016) compared the performance of the Sojourn method against their newly developed Sojourn-including Posture (SIP) method, which integrates information from the Actigraph and activPAL. Forty-nine adults wore the two monitors, with an Oxycon Mobile and direct observation used as criterion measures. Overall, the researchers found higher classification agreement for the SIP method (79%) than the Sojourn method (56%) based on direct observation. In addition, they found that the SIP method had lower mean absolute error than the Sojourn method at light, moderate, and vigorous intensity activity. A disadvantage of the SIP approach is that it requires two monitors to be worn simultaneously. Therefore, the Sojourn approach still merits further evaluation.

A logical step in advancing work on the Sojourn method is to evaluate the performance of this method under field conditions. Thus, the purposes of this study are: 1) to examine the performance of the Sojourn method for estimating energy expenditure and time spent at various intensities of activity relative to estimates from an established monitor-based measure (Sensewear Armband) and an established report-based measure (ACT24), and 2) to examine the context-related factors that are captured with the Sojourn method by comparing activity patterns with parallel data from the Sensewear Armband and ACT24. A detailed review of the literature will be provided in the next section to provide a justification for the measures and methods to be used in the study.

CHAPTER 2. REVIEW OF LITERATURE

The importance of physical activity has been noted for centuries. From ancient Chinese civilizations dating back 4,500 years, where organized exercise was first used to promote health, to Ancient Greece, where physician-teacher Hippocrates authored many treatises on physical activity and health in the 5th and 4th centuries BCE, physical activity has been a prominent force in ancient and contemporary culture (MacAuley, 1994). Despite this, it was not until the early 20th century when researchers began to examine the quantitative impact of physical activity on health. In 1939, O.F. Hedley analyzed the incidence of coronary heart disease in 5,000 men, finding that mortality among men in business and professional groups was considerably higher than men working in manual labor. At the time, this was not attributed to differences in physical activity (Paffenbarger, Blair, & Lee, 2001). However, Jerry Morris and colleagues (1953) discovered 14 years later that postal clerks, who spent most of the day sitting, had higher coronary heart disease rates than the postmen, who spent most of the day walking. This landmark study set the foundation for future investigations into the beneficial impact regular physical activity has on health (Paffenbarger, Blair, & Lee, 2001). Since then, regular physical activity has been shown to be effective in the long-term prevention of cardiovascular disease, diabetes, and certain cancers. Additionally, regular physical activity can improve energy, mood, endurance, energy balance, the ability to cope with stress, and the health of muscles, bones, and joints (CDC, 2015).

This evidence is paving the way for formalized U.S. Physical Activity Guidelines in 2018 and recommendations for updated guidelines have recently been released (Physical Activity Guidelines Committee, 2018). The goal is for adults to accumulate 150 minutes of moderate intensity physical activity, 75 minutes of vigorous intensity physical activity, or an equivalent combination of both across a week. Intensity of physical activity is most often classified into categories of sedentary, light, moderate, and vigorous based on ranges of MET values. Defined as the ratio of the rate of energy expended during activity to the rate of energy expended at rest, activity at 1.6-2.9 METs is light intensity, 3.0-5.9 METs is moderate intensity, and 6.0+ METs is vigorous intensity (ODPHP, 2017). The 2011 Compendium of Physical Activities offers a thorough list of activities and their associated absolute MET values (Ainsworth et al., 2011).

Despite the well-known benefits of physical activity, many do not meet these guidelines (Clarke, Norris, & Schiller, 2016). Per an early release of data from the 2016 National Health Interview Survey (NHIS), approximately 51.7% of U.S. adults aged 18 and over were found to have met the 2008 guidelines for aerobic activity (Clarke, Norris, & Schiller, 2016). This is a partial contributor to the high prevalence of obesity in society (Mitchell, Catenacci, Wyatt, & Hill, 2011). A recent surveillance study by Ogden Carroll, Fryar, and Flegal (2015) found that more than 36% of U.S. adults are obese, costing $147 billion annually and contributing to many chronic diseases.

While the promotion of physical activity has been a key goal there are related concerns about high levels of sedentary behavior in society. Sedentary behavior has been defined by the Sedentary Behavior Research Network as ‘any waking behavior characterized by an energy expenditure ≤1.5 metabolic equivalents (METs), while in a sitting, reclining, or lying posture’ (Tremblay et al., 2017). Activities such as television-watching and computer use are often used as proxies of sedentary behavior in scientific research, but the technical definition now captures both a postural component and an energy expenditure / MET component. A recent review of 22 studies by Harvey, Chastin, and Skelton (2015) found that adults aged 60 and over were sedentary for over 9 hours a day, according to objective measurement, equating to 65-80% of their waking day. Additionally, Matthews et al. (2008) found younger adults spending an average of 7.5 hours daily in sedentary pursuits. Extended periods of sedentary behavior may compromise metabolic health, increasing cardiovascular and all-cause mortality (Owen, Healy, Matthews, & Dunstan, 2010).

This literature review will summarize research on physical activity assessment technique. The first section will provide an overview of PA and SB assessment methods. The second section will provide a deeper coverage of issues with accelerometry-based measures since that will be the focus of the proposed study. The last section will specifically outline the gaps in the literature and provide a justification for the methods.

2.1 Overview of PA and SB Assessment Methods

There are a number of methods to assess PA and SB, but there are advantages and disadvantages to each. Two main categories are report-based measures, which include self-report questionnaires and activity diaries or logs, and monitor-based measures, which include heart-rate monitors, pedometers, and accelerometry-based activity monitors. The following sections provide an overview of each of these categories, with a focus on accelerometry-based activity monitors since they are a more commonly used tool in contemporary research.

2.1.1 Report-Based Methods

Report-based assessment methods for physical activity and sedentary behavior rely on the perceptions and recall of previously performed activities. Such methods are feasible and easy to use in large scale applications (Helmerhorst, Brage, Warren, Besson, & Ekelund, 2012). However, these survey-based tools are subject to the cognitive challenges associated with recall and suffer from inconsistent validity, limited by factors such as social desirability bias and the complexity of the questionnaire (Prince et al., 2008).

Among the most widely used report-based tools for assessing physical activity and sedentary behavior are self-report questionnaires (Castillo-Retamal & Hinckson, 2011). Questionnaires differ in several ways, depending on the target behavior, how the questionnaires are administered, and how data are reported (Prince et al., 2008). For example, many questionnaires that assess sedentary behavior target daily television viewing as a marker of overall sedentary behavior (Atkin et al., 2012). Similarly, many questionnaires targeting physical activity behavior capture time spent in moderate-to-vigorous intensity activity as a proxy of total physical activity, easily comparable to the guidelines offered by the USDHHS. A review of 65 studies examining 96 questionnaires by Helmerhorst, Brage, Warren, Besson, and Ekelund (2012) found the questionnaires to have median test-retest reliability correlation coefficients ranging from 0.62-0.76 and median validity coefficients ranging from 0.25-0.41. Thus, despite showing modest reliability, most questionnaires struggled to show acceptable validity. This is a challenge that researchers face when using self-report questionnaires and must be considered when choosing a measure to assess physical activity patterns.

Logs and activity diaries are one way to capture data about PA and SB. The use of these logs often requires participants to record their behavior and activities in real time. One commonly employed tool to assess physical activity, Bouchard’s Physical Activity Record (BAR), asks participants to record activity in 15-minute intervals over three days (Bouchard et al., 1983). Estimates from the Bouchard instrument are not highly accurate since the 15-minute blocks force participants to record a single activity for each block (Wickel, Welk, &

Eisenmann, 2006). In a review of physical activity and sedentary behavior assessment methods, Atkin et al. (2012) notes that diaries and logs may overcome the limitations of recall and bias associated with self-report questionnaires, if the target behavior is recorded in real time. However, the burden of such recording can be high, especially for those with cognitive dysfunction, which could hinder compliance and induce participant reactivity through extreme self-monitoring (Prince et al., 2008).

Another report-based approach is to ask participants to recall their previous activity patterns. A common format is the 7-day physical activity recall that asked participants to recall data over the past 7 days (Blair et el., 1985). Other common forms include the International Physical Activity Questionnaires (Craig et al., 2003) and the Modifiable Activity Questionnaire (Kriska et al., 1990). While many approaches have been used, there is some consensus that 24-hour recall formats provide some advantages over other longer or more generalized recall formats (Matthews, Moore, George, Sampson, & Bowles, 2012). There are a variety of forms and implementation methods for 24PAR instruments, but they generally guide participants through the previous 24 hours and asks them to recall activity in 4 distinct 6-hour time blocks: 12AM-6AM, 6AM-12PM, 12PM-6PM, and 6PM-12AM. An online version of the 24PAR called the ACT24 allows participants to select from 111 individual activities in 14 major categories, such as personal care or lawn and garden, and record how long they performed each activity as well as if they were sitting, standing, or both (if necessary). A recent study by Matthews et al. (2017) examined the utility of the ACT24 for estimating energy expenditure. The researchers studied 932 adults and compared various outcomes from the ACT24 with the reference measures of doubly-labeled water (DLW) and an activPAL monitor. The researchers found that ACT24 estimates for energy expenditure were within 3-10% of DLW energy expenditure measurements. In addition, ACT24 estimates for active and sedentary time were within 1-3% of estimations from the activPAL monitor.

An interview administered version of the 24PAR was used in a large study called the Physical Activity Measurement Survey (PAMS). This study involved a sample of 1,347 adults wearing a Sensewear Armband (SWA) for a 24-hour period and returning to complete a physical activity recall of the previous day's activities. Each participant performed two trials, and energy expenditure and MVPA values were averaged across the two days. Equivalence testing and calculations of mean absolute percent error were used to evaluate agreement between the SWA and ACT24. An evaluation of this data demonstrated good agreement between the two measures for capturing energy expenditure. MAPE values ranged from 10.3% (female) to 13.3% (males) for sex, 11.1% (50-71 years) to 15% (20-29 years) for age group, and 11.4%-12.1% for BMI. However, comparisons for MVPA were not found to be statistically equivalent. MAPE values ranged from 68.6% to 269.5% across all comparisons. Estimates of MVPA from the PAR tended to be underestimated for younger, less obese people but overestimated for older, more obese people. However, when using mean percent error, group-level estimates of EE and MVPA from the PAR were within 10% of SWA values (Welk et al., 2014).

In an additional analysis from the same PAMS project, Kim and Welk (2017) aimed to determine the validity of the PAR relative to the SWA for assessing total sedentary time. When not including reported or recorded sleep time in the estimate of total sedentary time, analysis revealed lack of equivalence between the recall tool (90% CI: 443.0 and

457.6 min/day) and the SWA (equivalence zone: 580.7 and 709.8 min/day). However, error was smaller between the measures for from those who were minimally or extremely active.

These findings suggest that estimates of EE, MVPA, and sedentary time from 24PAR instruments yielded similar estimates as those provided by various wearable monitors (for group level comparisons). The strength of these 24PAR formats seems to be in its estimation of energy expenditure, with studies finding less than 15% error (Welk et al., 2014; Matthews, 2017). The findings from these studies support the use of these 24PAR formats for EE estimation, but error for estimates of MVPA and SB is likely high for individual applications. However, calibration methods offer promise for improving the accuracy of both PA and SB. The proposed study will utilize the online version of the Act24 to provide context about PA and SB behavior.

2.1.2 Monitor-Based Methods

Rather than relying on individuals to recall or record their own activity, the goal of monitor-based measures is to objectively assess physical activity or sedentary behavior. Generally, such methods are not subject to errors in recall or bias from the participant. However, monitor-based methods can be difficult to employ in large-scale surveillance studies due to financial cost and the burden of processing vast amounts of data, and the field lacks standards of practice as it relates to monitor calibration and validation (Freedson, Bowles, Troiano, & Haskell, 2012). The category of monitor-based methods includes pedometers, heart rate monitors and GPS methods, but the focus in this review is on accelerometry-based activity monitors since they are viewed as the most promising and widely used method.

A wide range of accelerometry-based activity monitors (hereafter referred to as accelerometers) are available for use in research and consumer applications. They vary depending on the model, location (hip, thigh, wrist, arm, etc.), and the number of axes (uniaxial vs. triaxial) from which the device detects acceleration. Triaxial accelerometers operate by detecting acceleration across three planes, which are the anteroposterior, mediolateral, and vertical planes. The signal detected from this acceleration has historically been summarized into “counts” over a user-defined time, called an epoch. A higher count value implies a greater intensity or volume of PA, but the interpretation of counts has been challenging. (Vanhelst et al., 2012). Researchers have developed methods to convert counts into a number of different, more practical outcomes, such as energy expenditure and MVPA, through the application of regression models and other calibration methods. Generally, the strengths of accelerometers include the ability accurately categorize activity by intensity, capture large amounts of data, and monitor activity minute-by-minute. However, they can be expensive to administer in large-scale applications, and large amounts of data can be unwieldy, requiring technical expertise and making data collection and analysis burdensome. In addition, accelerometers provide no information about the context of behavior. Thus, report-based measures are often in tandem with monitor-based methods to provide contextual information as it relates to behavior and activity patterns.

2.2 Accelerometers: A Closer Look

The rise of accelerometry as a method of assessing physical activity and sedentary behavior has led to a vast literature on a variety of monitors. The following section reviews literature on various monitors and primary issues associated with each. The review begins with coverage of the Actigraph monitor which is the most commonly used device to provide an overall view of the literature. Other devices and methods are then introduced since they provide alternatives to the widely used Actigraph.

2.2.1 Actigraph

The Actigraph (Pensacola, FL, USA) is perhaps the most widely-studied monitor in the field, accounting for more than 50% of published studies (Migueles et al., 2017). Early studies made use of the uniaxial Actigraph accelerometer, such as the model 7164 or GT1M, which only detected vertical acceleration. Their widespread use in the 1990's and early 2000's led to the development of cut-points and algorithms that assessed sedentary behavior, intensity of activity, and energy expenditure from vertical axis accelerations (Sasaki, John, & Freedson, 2011; Migueles et al., 2017). The original Freedson, Melanson, and Sirard (1998) cut points for uniaxial applications were widely used in the field. They defined light activity as anything less than or equal to 1,951 counts per minute, moderate intensity physical activity from 1,952-5,724 counts, vigorous intensity physical activity at 5,725-9,498 counts, and very vigorous physical activity at 9,499 counts or greater.

Many older studies have evaluated the use of uniaxial Actigraph accelerometers in physical activity research. An early study by Janz (1994) compared counts from an Actigraph uniaxial accelerometer with heart rate in 31 participants. The monitors were worn for three consecutive days for 12 hours while youth engaged in variety of free-living activities. Findings indicated moderate to high validity correlation coefficients (r = 0.50-0.74) and a high correlation between counts from the Actigraph and heart rate (r = 0.69). In a later study by Trost et al. (1998), 30 youth wore one Actigraph 7164 on each hip during three 5-minute treadmill bouts at 3, 4, and 6 mph with VO_2 being monitored by an automated system. Using an energy expenditure prediction equation, monitor estimations were compared to energy expenditure estimations from VO_2 formulas. Findings from this laboratory-controlled study indicated that activity counts from each monitor were strongly correlated with energy expenditure (*r* = 0.86 and 0.87). In addition, the energy expenditure prediction equation was able to predict mean energy expenditure within 0.01 kilocalories/min. Despite these findings, the limitations of uniaxial accelerometry are well known. A review by Corder, Brage, and Ekelund (2007) noted that the main limitation of uniaxial accelerometry lies in its inaccuracy for capturing non-ambulatory movement, such as cycling. If one's movement is largely ambulatory, this issue is lessened. However, uniaxial accelerometers have also been found to struggle at higher speeds of ambulatory activity. Counts per minute (cpm) have been found to increase linearly during activity until approximately 10,000 cpm in adults, where they plateau. As a result, no significant difference can be found between counts at running speeds of 20-26 km/h and counts at a walking speed of 6 km/h (Corder, Brage, & Ekelund, 2007).

In 2009, Actigraph released the triaxial GT3X activity monitor, which utilizes a tri-axial accelerometer that captured movement across three different planes (anteroposterior, mediolateral, and vertical) (Sasaki, John, & Freedson, 2011). The change led to gradual shifts away from uniaxial monitors to triaxial accelerometers. One such study by Kelly et al. (2013) observed the correlation and agreement between counts from the Actigraph GT1M (uniaxial) and GT3X (triaxial) with VO_2. Forty-two participants wore the GT1M and GT3X on their right hip during laboratory-controlled treadmill exercises at three different speeds: slow walking at 4.8 km/h, fast walking at 6.4 km/h, and running at 9.7 km/h. The correlations between counts from the GT1M and the GT3X and VO_2 were high (*r* = 0.881 and 0.810).

Despite the widespread use, there are many limitations with the use of the Actigraph. Specifically, a review by Atkin et al. (2012) noted that many studies have found that the

Actigraph has difficulty in distinguishing changes in posture. Thus, it may categorize some activity as sedentary (sitting or lying down) when a participant may be engaged in light activity (standing). However, newer models include an inclinometer, which could be a potential remedy for this issue. The growth of the accelerometry field has led to the development of many data processing algorithms for Actigraph data and there is little consensus about which methods to use. The number of data processing decisions make it difficult for researchers to make appropriate choices (Migueles et al., 2017). Moreover, the wide variability in outcomes from different methods limits comparability in research using different approaches. The focus of the study is on a more promising method for using the Actigraph, but these will be discussed in the next section.

2.2.2 Sensewear Armband

The Sensewear monitor was introduced as an innovative alternative to traditional accelerometry-based monitor. The key innovations were the use of multiple sensors and the incorporation of pattern recognition technology that were used to collectively estimate energy expenditure. Its associated software program allows researchers to analyze data in terms of raw counts, energy expenditure in joules or kilocalories, METs, steps taken, and activity duration

The monitor is worn on the arm and uses a motion sensor to detect acceleration as well as heat-related sensors, such as heat flux, galvanic skin response, skin temperature, and body temperature, to measure energy expenditure. Using this dual measurement strategy, the Sensewear Armband has been found to be more sensitive to measuring energy expended during complex and non-ambulatory activities. In a study of 30 adults over 2 weeks, Johannsen et al. (2010) tested agreement between two Sensewear Armbands and goldstandard method of doubly-labeled water (DLW) for energy expenditure. Participants wore the Sensewear Pro3 Armband (SWA) and the Sensewear Mini Armband (Mini) for 14 consecutive days, including while sleeping. Findings indicated that absolute error rates were similar for the two monitors (SWA = 8.1%, Mini = 8.3%). Additional intraclass correlation (ICC) analysis revealed significant agreements between the SWA and DLW estimates of energy expenditure (ICC = 0.80).

In a similar study by Mackey et al. (2011), the utility of the Sensewear Armband for estimating energy expenditure was assessed using two different versions of the Sensewear program. Nineteen older adults wore a Sensewear Armband for a mean of 12.5 days, including while sleeping. During this period, energy expenditure was assessed using doubly labeled water. Findings on total energy expenditure indicated no significant difference in mean estimates from doubly labeled water (2,040 ± 472 kcal/day), SWA 6.1 software version (2,012 ± 497 kcal/day) or SWA 5.1 software version (2,066 ± 474 kcal/day). In addition, individual estimates from each method were highly correlated with one another (SWA 6.1 r = .893; SWA 5.1 r = .901) and demonstrated strong agreement (SWA 6.1 ICC = .896; SWA 5.1 ICC = .904). Bland-Altman plots identified no systematic bias for TEE or AEE. Agreement analysis using Bland-Altman plots also identified no systematic bias for estimates of total energy expenditure.

A key advantage of the Sensewear technology is that it was iteratively improved over time as the pattern recognition algorithms were improved. For example, a study by Lee, Kim, Bai, Gaesser, & Welk (2016) evaluated validity of different Sensewear software (algorithms v5.2 vs. algorithm v2.2) for estimating energy expenditure (EE) in children. Forty-five children aged 7–13 years performed 12 randomly assigned activities while wearing a

Sensewear with simultaneous monitoring via portable calorimetry (IC). Each activity lasted 5 minutes, with a 1-minute break between activities. Analyses revealed smaller errors for the newer v5.2 algorithms (0.25 ± 0.09 kcal min^{-1}) than the older v2.2 algorithms (1.04 ± 0.09 kcal min^{-1}). The mean absolute percent error (MAPE) was $17.0 \pm 12.1\%$ for Sensewear v5.2 algorithm and $31.4 \pm 11.1\%$ for Sensewear v2.2 algorithm. The v5.2 algorithms yielded non-significant ($p > 0.5$) differences in EE estimates for most of the walking related activities as well as for stationary cycling at moderate intensity (MAPE = 14.5%).

In a similar study by Lopez, Brond, Andersen, Dencker, and Arvidsson (2017) the accuracy of the Sensewear algorithm v5.2 was tested relative to a previous version v2.2. These algorithms were evaluated for estimates of energy expenditure during various activities, such as sitting, walking, running, and biking relative to indirect calorimetry. In total, 35 children, 31 adolescents, and 36 adults participated. Analyses revealed mean absolute percent error for all activities with Sensewear v5.2 were 24% for children, 23% in adolescents, and 20% in adults. Comparatively, mean absolute percent error for corresponding activities with Sensewear v2.2 were 37% for children, 26% for adolescents, and 25% in adults, indicating improvement in all age groups for the Sensewear v5.2 algorithm.

These studies are a few examples of comparative studies that have tested differences in algorithms. The Sensewear technology has been considered by many researchers to provide the strongest approach for assessment of physical activity and sedentary behavior but BodyMedia was bought by Jawbone in 2013 and the monitor line was discontinued. While the monitor is no longer commercially available it still provides a useful field-criterion of accurately assessing energy expenditure. It is used in the present study as a comparison measure.

2.3 Overview of Sojourn Method

A key issue with accelerometry lies in the way data are expressed. Raw accelerometer data are expressed in counts per unit of time. As a unit, counts mean little on their own as it relates to physiology and when translating research into practical outcomes. Thus, there has been a need for calibration methods that translate counts into more meaningful units. As a result, researchers have developed regression formulas to derive MVPA and energy expenditure from accelerometer data in the form of cut-points (Miller, Strath, Swartz, & Cashin, 2010). There are a wide variety of cut-points to choose from when translating accelerometer data. For example, laboratory-based cut-points for youth hip-worn data developed by Freedson, Pober, and Janz (2005) categorize sedentary behavior as anything less than or equal to 100 counts per minute, light physical activity at 101-2,219 counts, moderate intensity physical activity at 2,220-4,135 counts, and vigorous intensity physical activity at 4,136 counts and greater. Another cut-point method for youth developed by Treuth et al. (2004) define sedentary behavior as anything less than or equal to 100 counts per 30 seconds, light physical activity at 101-2,999 counts, moderate intensity physical activity at 3,000-5,200 counts, and vigorous intensity physical activity at 5,201 counts and greater. For adults, early cut-points from Freedson, Melanson, and Sirard (1998) defined light activity as anything less than or equal to 1,951 counts per minute, moderate intensity physical activity from 1,952-5,724 counts, vigorous intensity physical activity at 5,725-9,498 counts, and very vigorous physical activity at 9,499 counts or greater.

Such cut-points have been tested in a variety of settings and applications in an effort to derive the most accurate methods. The wide variety of cut-points that are available, however, draw caution from researchers (Prince et al., 2015). Many cut-point equations were derived to broadly categorize the intensity of physical activity (Bassett, Rowlands, & Trost, 2012). However, certain considerations must be made for the population of study (age, health, movement ability, etc.), the indicators of interest (sedentary time, energy expenditure, etc.), and the type of movement being studied (non-ambulatory vs. ambulatory). Failure to do so may have a wider impact on physical activity surveillance research, particularly with how well the population is judged to meet physical activity guidelines. For example, in a study of nine cut-points in a national sample of over six thousand, Watson, Carlson, Carroll, and Fulton (2014) found that the prevalence of those in the sample meeting 2008 Physical Activity Guidelines was considerably lower for cut-points derived from ambulatory methods (median=11.5%, range=6.3%–27.4%) compared to cut-points derived from lifestyle protocols (median=77.2%, range=60.6%–98.3%). Cut-point choice can impact multiple arms of research and is critical to the field of physical activity measurement.

Recently, Lyden, Kozey-Keadle, Staudenmeyer, and Freedson (2014) developed the Sojourn method for calibrating Actigraph data. It is a hybrid machine learning technique that combines artificial neural networks with decision tree analysis. This free-living method uses counts per second from signals in three different axes from the hip-mounted triaxial Actigraph accelerometer. First, the method identifies when bouts of activity and inactivity start and stop by identifying instances of rapid acceleration or deceleration. These instances are defined as any movement that produces a vertical acceleration signal equal to or greater than 15 counts per second. Next, the method determines whether these bouts are activity or inactivity. The method identifies activity out of a series of zero (inactivity) and nonzero (activity) counts. Thus, activity is distinguished from inactivity using the percentage of nonzero counts from the vertical axis in the interval. Vertical axis nonzero counts greater than or equal to 70% constitutes activity. Inactivity is categorized into four types: sitting inactivity type 1 (nonzero counts ≤5%), sitting inactivity type 2 (>5%), standing inactivity type 3 (≤12%), and standing inactivity type 4 (>12%). Once bouts of activity and inactivity are determined, non-PA MET values are assigned to inactivity bouts and MET values for activity bouts are estimated using an artificial neural network. The non-PA MET values are based on the Compendium of Physical Activities, with inactivity type 1 equal to 1 MET, inactivity type 2 equal to 1.2 METs, inactivity type 3 equal to 1.5 METs, and inactivity type 4 equal to 1.7 METs. MET values are estimated for activity based on bout length. Less than 120 seconds, the artificial neural network is applied to the entire bout. Longer than 120 seconds, and the artificial neural network segments the bout into 40 second intervals, estimating a MET value for each (Lyden, Kozey-Keadle, Staudenmayer, & Freedson, 2014).

The Sojourn method may have advantages over previous techniques. One such advantage is how the signal is segmented. Previous approaches function by integrating acceleration signals from the various planes over a user-defined time sampling interval, called an epoch (Gabriel et al., 2010). Epochs can differ in length (1 second to 1 minute) depending on memory capacity of the accelerometer. At the end of each epoch, counts are stored in the monitor’s memory (Gabriel et al., 2010). However, the performance of this sliding window model declines when applied to free-living activities that are unplanned and of different length (Lyden, Kozey-Keadle, Staudenmayer, & Freedson, 2014; Crouter, Kuffel, Haas, Frongillo, & Bassett, 2010). The Sojourn method avoids this issue by using nonfixed, activity-defined windows to segment the signals. As a result, the use of this method may lead to greater accuracy for estimating outcomes related to time spent in MVPA and energy expenditure.

Despite the promise of the Sojourn method, few studies have compared it to established measures. Recently, Ellingson et al. (2016) compared the performance of the Sojourn method against their newly developed Sojourn-including Posture (SIP) method, which integrates information from the Actigraph and activPAL. Forty-nine adults completed five-minute bouts of 15 activities ranging in intensity from sedentary (e.g. sitting reading a book) to vigorous (running at 5.5 mph) while wearing an Actigraph on the hip, activPAL on the thigh, and an Oxycon Mobile. Energy expenditure and activity intensity estimates were gathered from both the Sojourn method (Lyden, Kozey-Keadle, Staudenmayer, & Freedson, 2014) and their new SIP method. Findings indicated that the SIP method had a higher overall classification agreement (79%) than the Sojourn method (56%) based on direct observation. When compared to the Oxycon Mobile, energy expenditure estimates derived from the SIP method had lower mean absolute error than the Sojourn method for light-intensity (0.21 vs 0.27), moderate-intensity (0.33 vs 0.42), and vigorous-intensity (0.16 vs 0.35) activities.

Another study by Kim, Barry, and Kang (2015) aimed to validate the Actigraph GT3X and the activPAL™ for assessment of sedentary behavior against video observation. Eleven participants wore each monitor as well as a wearable camera over a 6-hour period, in which the researchers were attempting to capture freely-occurring sedentary behavior. Time spent in sedentary behavior was estimated from the Actigraph using traditional activity count thresholds (<150 cpm) and the Sojourn method. Both the Actigraph and the activPAL™ were compared to video observation. Findings indicated that the activPAL™ assessed sedentary behavior most accurately. However, the Sojourn method was more accurate than traditional cut-point methods in identifying sedentary behavior when bouts are 15 minutes or less.

A logical step in advancing accelerometry research is to compare new calibration methods for predicting time spent at different intensities of activity and energy expenditure. The Sojourn method may have advantages over previous approaches, but fewer studies have compared it to other, more-established measures. Thus, there is a need to test the performance of the Sojourn method against various comparison measures for various outcomes, such as time spent in MVPA and energy expenditure.

CHAPTER 3. METHODS

Overview of Study and Design

The study used data obtained through a large (ongoing) field-based evaluation of activity monitors conducted in the Physical Activity and Health Promotion Lab at Iowa State University (www.physicalactivitylab.org). The study used a convenience sampling plan to recruit a sample of 105 adults aged 18-60 from the state of Iowa. Participants were recruited through various avenues. Flyers detailing the study were posted around Iowa State University's campus. To increase representation of older adults, flyers were additionally posted around various high-traffic areas in the community of Ames, Iowa. Compensation for participation in the study was $25. Participants were required to be between the ages of 18-60 and able to move about on their own. Exclusion criteria included individuals with metal allergies (to avoid reaction to the monitors) and those with any medical injury or condition that prevented movement by walking. Recruitment and study procedures were approved by the local Institutional Review Board and written consent was obtained from all participants.

Instruments

Sensewear Core Armband (SWA)

The SWA (BodyMedia, Inc., Pittsburgh, PA, USA) is a multi-sensor activity monitor worn on the back of an upper arm. Through a combination of multiple sensors, including the detection of heat flux, galvanic skin response, skin temperature, and near body temperature, and a tri-axial accelerometer, the SWA estimates time spent at various intensities of activity, total energy expenditure, active energy expenditure, average METs, physical activity duration, and number of steps. SWA data was processed using the latest version of software v8.0 (algorithm v5.2).

Actigraph wGT3X-BT

The wGT3X-BT (Actigraph Corporation, Pensacola, FL, USA) is an activity monitor that can be worn in multiple locations, including the wrist, waist, ankle, and thigh. It includes a micro-electro-mechanical system (MEMS) based accelerometer, an ambient light sensor, and an inclinometer. Users can select sample rate (30Hz-100Hz) pre-data collection and filtering size (1-60 second epochs) post-data processing. In this study, the wGT3X-BT was initialized at 100Hz, and 1-second epochs were used as part of the Sojourn method analysis. Acceleration data are gathered across three axes (vertical, horizontal, perpendicular) and processed using software v6.11.6.

ACT24

The ACT24 is an online self-report recall tool designed to capture activity over a 24-hour period. Researchers in the lab initialized the instrument for use, but the tool itself is self-guided. It guides participants through the previous 24 hours and asks participants to recall activity in 4 distinct 6-hour time blocks. Participant select from 111 individual activities in 14 major categories, such as personal care or lawn and garden, and record how long they performed each activity, the intensity (light, moderate, vigorous), as well as if they were sitting, standing, or both (if necessary). Only activities lasting 5 minutes or more were recorded. The data are stored in a secure, online server, and were downloaded post-visit.

Data Collection

Data were collected by an experienced and trained research group at Iowa State University. All researchers completed relevant human subjects research training. In addition, researchers were extensively trained on protocol and data processing before being allowed to gather data. Data were collected using a standardized procedure. Participants were asked to complete two appointments with researchers, separated by a 24-hour period of activity monitoring.

First Appointment

To begin the first meeting, participants gave written consent to participate in the study. A paper-administered demographic survey was administered to identify various participant characteristics and determine eligibility for participation. Questions included age, sex, ethnicity, employment, income, marital status, education, previous use of a physical activity monitor, perception of health status, and physical activity status. Following this process, anthropometric measurements were completed. This was completed in a private room with a maximum of two researchers present.

At the end of the first visit, the activity monitors were given to the participants to wear for a full 24-hour period. Written and verbal instructions were provided as to how to wear the monitors. Participants were instructed to wear 3 Actigraph accelerometers (1 on each wrist, 1 on waist), 1 Sensewear Armband (back of arm), 1 activPAL™ (thigh), 1 GENEactiv (wrist), and 1 Axivity (wrist) for the full 24-hour period on the day following the appointment and were provided with a log to record duration and type of activity performed if the monitor was removed. Though seven monitors were included within the protocol as part of the larger study, this study only made use of the hip-worn Actigraph and the Sensewear Armband. The first meeting lasted approximately 20 minutes.

Second Appointment

On the day following the data collection period (two days following the first appointment), participants returned to the lab. The monitors along with the log were returned to researchers at this time. During the second meeting, participants completed the ACT24.

The online self-report tool was set up by members of the research team. Researchers offered direction as to how to complete the assessment, but the tool itself was self-administered. It required participants to record activity over the full 1,440 minutes when the monitors were worn. The second and final meeting lasted approximately 30 minutes.

Data Processing

Data from the SWA and ACT24 were processed to facilitate direct comparisons of time spent in various intensities of activity with data from the Actigraph using the Sojourn calibration method. SWA data were processed using the Sensewear Professional software. Individual raw Sensewear files were exported into .csv or .xlsx file types. Individual files were then compiled into one summary dataset to facilitate dataset merging with results from other methods. Indicators of physical activity included energy expenditure and time spent at various intensities of activity (minutes). Allocations for minutes spent in moderate physical activity (3-6 METs), vigorous physical activity (6-9 METs), and very vigorous physical activity (> 9 METs) were combined to obtain an indicator of time spent in MVPA. Time spent in vigorous intensity activity was obtained by combining the categories of vigorous and very vigorous activity. In addition, sedentary time values included sleep to facilitate analysis with comparison methods.

Actigraph data were processed through a combination of the ActiLife and R programming software programs. Data were initially processed through Actigraph’s parent program, ActiLife. Using this program, individual 1-second hip Actigraph files were downloaded stored on a secure drive. Subsequently, syntax written in the statistical computing program R searched the drive where the files were stored and ran the Sojourn method on individual hip Actigraph files. Briefly, the Sojourn calibration method identified when bouts of activity and inactivity start and stop by identifying instances of rapid acceleration or deceleration (vertical acceleration signals equal to or greater than 15 counts per second). Bouts of inactivity were categorized into four types with assigned MET values: inactivity type 1 equal to 1 MET (sitting or lying still), inactivity type 2 equal to 1.2 METs (sitting with little movement), inactivity type 3 equal to 1.5 METs (standing still), and inactivity type 4 equal to 1.7 METs (standing with little movement). MET values for activity bouts were estimated using an artificial neural network. Raw Sojourn files (second-by-second) were saved along with a compressed file containing summary indicators of energy expenditure and time spent at various intensities of activity.

Data from the ACT24 were downloaded from the ACT24 Researcher Site. A summary file containing data from all participants was downloaded as a .7z extension, which is a type of compressed folder that requires 7zip software to open. The .csv file contained within this extension revealed data on several indicators from each participant. Time spent at each intensity level is directly given. Estimates from time spent in moderate intensity physical activity and vigorous intensity physical activity were combined to obtain an indicator of MVPA. Energy expenditure is not directly measured by the ACT24, but METs are. Thus, METs were multiplied by time to obtain an indicator of energy expenditure.

Once data related to relevant indicators were compiled for all 3 methods, each dataset was merged within R to create a “master dataset” containing all data. Data were merged by participant ID, and correct matching was ensured by an additional merging of datasets within Excel. The resulting file was stored in R as well as exported to Excel to facilitate data analysis.

Data Analysis

Several analyses were run to assess the performance of the Sojourn method for capturing energy expenditure and time spent at various intensities of activity against various comparison measures, as well as examine the context-related factors captured by the Sojourn method. The Sensewear Armband was used as the established comparison measure, while the ACT24 was used as a comparison tool for various physical activity indicators and contextual activity information.

Descriptive Analyses

Participant demographics were summarized to document the characteristics of the sample since this can influence findings and observed relationships. In addition, stacked bar graphs were used to look at time allocations obtained from each of the three methods. In a similar vein to the participant demographic table, these graphs were used to generally summarize participant physical activity and assist in the interpretation of the findings.

Method Agreement Analyses

Comparison and error analyses were conducted to assess method agreement, particularly between the Sojourn method and the two comparison methods. Correlational analyses were used as surface-level methods to assess the relationship between indicators from the Sojourn method with indicators from the Sensewear Armband and the ACT24. Specifically, Pearson's r was used to denote the relationship between the methods for assessing total daily energy expenditure and time spent at various intensities of activity. Correlation estimates were gathered for all indicators and method (Sojourns and SWA, Sojourns and ACT24, SWA and ACT24).

The magnitude or error was quantified using established indicators to assess the performance of the Sojourn method for capturing energy expenditure and time spent at various intensities of activity. Specifically, mean percent error (MPE) was calculated for the Sojourn method, using the SWA as the established reference measure. Individual error was calculated for each participant over the day and then averaged to obtain an overall indicator of MPE, expressed as a percentage. In addition to MPE, mean absolute percent error (MAPE) was also used to characterize total error irrespective of the cancellation of overestimation and underestimation that could be seen in the calculation of MPE. Individual MAPE of estimates from the Sojourn method were calculated for each case relative to estimates from the SWA, and then averaged to obtain a total indicator of MAPE. Total values for both MPE and MAPE were calculated for each indicator of interest. Values were calculated within R in addition to Excel.

Error Visualization

To visualize error between the methods, Bland-Altman plots were also used. Bland-Altman plot analysis is a method to quantify agreement between two quantitative measurements by establishing limits of agreement, which are calculated by using the mean and the standard deviation of the differences between two measurements. Represented in graphical form, the difference between the two measures (y-axis) is plotted against the mean of the two measurements (x-axis) (Giavarina, 2015). The primary focus was to examine error between the Sojourn method and the SWA; thus, the two methods were plotted against one another using Bland-Altman plots to visualize magnitude and direction of error.

CHAPTER 4. RESULTS

Out of 105 total cases, 20 had missing data due to monitor malfunction, failure to properly initialize monitors, or error in the self-report system. Thus, 85 cases (54 female) with a full set of observations were used in the analyses. Participant demographics are summarized in Table 1. Mean age was 26 years, with a majority (80%) aged between 18 and 30, and were mostly White, not Hispanic (76%). Most had at least some higher education (87%), with an overrepresentation (29.4%) of those with an advanced degree (Masters and further), compared to the general population. Most answers on the demographic survey likely indicated student status, with a majority making less than $25,000 (58.8%) annually and working part time (42.4%). Though most reported currently being physically active (90.6%) and engaging in regular (as defined by guidelines) physical activity (78.8%), over one-third (34.1%) of participants reported never having used an activity monitor before. Anthropometric measurements revealed that a majority of participants (51.8%) were of normal weight according to current BMI standards (mean = 25.3 kg/m2, SD ± 5.1).

Table 2 and Figure 1 provide a breakdown of the average energy expenditure and activity profile as detailed by each of the three methods. Sample averages along with a breakdown by sex are provided. On average, the SWA (2869.0 kcal ± 788.7) recorded the greatest amount of expended energy over the 24-hour period compared to the ACT24 (2802.0 kcal ± 667.2) and the Sojourn method (2665.2 ± 745.0). This pattern was evident for both male subgroup (Sojourns: 3004.9 kcal ± 862.2, SWA: 3282.5 kcal ± 526.4, ACT24: 3245.6 kcal ± 874.4) and female subgroup (Sojourns: 2470.2 kcal ± 593.8, SWA: 2526.1 kcal ± 579.8, ACT24: 2652.9 kcal ± 650.1).

Participants spent the most time in sedentary pursuits. On average, participants over-reported their own sedentary time (1128.0 min ± 141.8) compared to the Sojourns (1015.3 min ± 145.1) and the SWA (989.1 min ± 181.9). Comparatively, the Sojourns recorded the most light activity (336.9 min ± 140.5) compared to the SWA (258.0 min ± 94.2) and the ACT24 (202.6 min ± 141.8). Despite the majority of time being spent in sedentary or light-intensity pursuits, a considerable amount of MVPA was detected by all three methods. On average, the greatest amount of MVPA was detected by the SWA (141.6 min ± 89.1) compared to the ACT24 (109.4 min ± 104.5) and the Sojourns (87.9 min ± 43.2). The most commonly reported activities were in the categories of transportation and structured exercise, as indicated by the ACT24. Popular forms of active transportation included walking (to and from classes, residence, bus stops, etc.) and biking, while many also engaged in structured cardio or weight-lifting regimens.

Categorization of the average activity profile by sex revealed similar trends but some gender differences were apparent (See Table 2). A total of 31 male cases and 54 female cases were used in the final analyses. Estimates of sedentary time were consistently higher (on average) for females (Sojourns: 1025.4 min ± 128.3, SWA: 1030.9 min ± 161.4, ACT24: 1134.1 ± 142.7) compared to males (Sojourns: 987.0 min ± 179.1, SWA: 916.2 min ± 194.8, ACT24: 1117.5 ± 163.0). The opposite pattern was found for light intensity activity, with the male subgroup averaging more time in activity of this intensity (Sojourns: 358.4 min ± 163.4, SWA: 283.4 min ± 84.1, ACT24: 214.7 min ± 154.6) compared to the female subgroup (Sojourns: 329.5 min ± 130.1, SWA: 243.5 min ± 97.4, ACT24: 195.7 min ± 134.9).

Estimates of MVPA from the methods also varied by sex. Males engaged in more MVPA with the Sojourn (94.6 min ± 37.4) and SWA (157.4 min ± 70.7) method compared to females (Sojourn: 85.1 min ± 46.9, SWA: 132.6 ± 97.6). However, the female subgroup self-reported more MVPA (110.3 min ± 115.8) on the ACT24 compared to the male subgroup (107.9 min ± 82.9), albeit non-significantly. In most cases, standard deviations indicated more variability in the average male activity profile for both sedentary time and light intensity activity, while the trend was reversed for MVPA, with more variability in the average female activity profile.

Surface-level analysis of method performance and agreement was achieved through correlational analyses. These analyses were performed for each indicator, including time spent in various intensities of activity and energy expenditure. Correlational analysis of the former indicated that most matching comparisons had at least a moderate association (see Tables 3, 4, and 5). When comparing the Sojourn method and the SWA, the strongest correlation was found for moderate intensity activity (r = 0.60, $p < .01$). This was followed by vigorous intensity activity (r = 0.46, $p < .01$), MVPA (r = 0.45, $p < .01$), and sedentary time (r = 0.38, $p < .01$). The correlation between the two methods for light intensity activity was weak and non-significant (r = 0.05). Comparatively, correlational analyses between the Sojourn method and ACT24 for the same indicator revealed a slightly different trend. The correlation between the two methods for moderate intensity activity was weak and non-significant (r = 0.17). However, moderate-strength relationships were found for each of the other activity intensities, including sedentary time (r = 0.52, $p < .01$), vigorous intensity (r = 0.46, $p < .01$), light intensity (r = 0.38, $p < .01$), and MVPA (r = 0.36, $p < .01$).

Additional correlational analyses compared the SWA and the ACT24. The strongest relationship was found for vigorous intensity activity (r = 0.69, $p < .01$), followed by sedentary time (r = 0.31, $p < .01$) and MVPA (r = 0.26, $p < .05$). Correlations for both light (r

= 0.12) and moderate intensity activity (r = 0.11) were weak and non-significant. Correlations between the methods for non-matching activity intensities (e.g. sedentary time and light intensity activity) were also computed, and can be found in Tables 3, 4, and 5.

Correlational analyses between the methods were also conducted for energy expenditure and can be found in Table 6. The strongest correlation was found between the Sojourns and ACT24 (r = 0.91, $p < .01$). Moderate strength relationships were also found between the Sojourns and the SWA (r = 0.65, $p < .01$) as well as the SWA and ACT24 (r = 0.59, $p < .01$).

Further analysis of method performance and agreement for each indicator was achieved through the computation of mean percent error (MPE) and mean absolute percent error (MAPE). These analyses as they relate to energy expenditure are summarized in Table 7. Analysis of MPE provides an indicator of the average error for groups of individuals. The Sojourn method was found to underestimate EE relative to the SWA (3.6%) and the Act24 (6.6%). The Act24 also underestimated EE relative to the SWA (4.4%). The calculations of MAPE provide a more robust indicator of the estimated error for individuals. The Sojourn method yielded a MAPE of 16.1% relative to the SWA and a MAPE of 9.5% relative to the Act24. The MAPE between ACT24 and SWA, for comparison, was 17.1%.

When computing MPE for activity intensity, the size of error varied between different intensities, as found in Table 7. Collectively, the smallest error between the methods was found for sedentary time and MVPA. The Sojourn overestimated sedentary time by 5.6% when compared to the SWA and underestimated by 12.5% when compared to the ACT24, respectively. When compared to the ACT24, the SWA underestimated sedentary time by 18.4%, on average. Mean percent error analyses of the methods for MVPA revealed a 17.6%

underestimation by the Sojourns compared to the SWA, a 30.0% underestimation by the Sojourns when compared to the ACT24, and 6.7% underestimation by the SWA when compared to the ACT24. Larger mean percent error values were found for light and moderate intensity activity. On average, the Sojourns overestimated light intensity by 63.6% when compared to the SWA and by 33.9% when compared to the ACT24. In addition, the SWA overestimated light intensity activity by 9.6% when compared to the ACT24. Conversely, the Sojourns underestimated moderate intensity activity by 37.5% when compared to the SWA and by 83.4% when compared to the ACT24, while the SWA underestimated moderate activity by 2.1% when compared to the ACT24. Given the small amount of vigorous intensity activity within the sample, MPE analysis was not reported.

Similarly, analysis of MAPE between the methods for activity intensity varied between intensities and are summarized in Table 8. The smallest MAPE values were found for sedentary time. Total error between the Sojourns and the SWA amounted to 17.1% for sedentary time, compared to 15.6% between the Sojourns and ACT24 as well as 23.7% between the SWA and ACT24. Larger error was found for MVPA, light intensity, and moderate intensity activity. Total error between the Sojourns and SWA for MVPA was 45.0% compared to 77.2% between the Sojourns and ACT24 as well as 84.8% between the SWA and ACT24. Mean absolute percent error values for light intensity activity were 82.8%, 50.0%, and 55.1% for the Sojourns and SWA, Sojourns and ACT24, and SWA and ACT24, respectively. The largest MAPE values were found for moderate intensity activity. For this intensity, total error was 59.2% between the Sojourns and SWA, 134.8% between the Sojourns and ACT24, and 93.9% between the SWA and ACT24. Like MPE analyses, MAPE analysis for time spent vigorous was unsuitable due to the little activity at this intensity.

To visualize error quantitatively and qualitatively, Bland-Altman plots were created for each indicator and are provided in Figures 2, 3 and 4. Plots for energy expenditure can be found in Figure 2. When comparing the Sojourns and the SWA, the Sojourns underestimated energy expenditure by 137 kilocalories, on average. Similarly, the Sojourns underestimated energy by 204 kilocalories when compared to the ACT24. When plotting energy expenditure as found by the SWA and ACT24, the ACT24 overestimated energy expenditure by an average of 67 kilocalories. Limits of agreement were narrower for the ACT24 and Sojourn plot than for the other comparisons, indicating greater consistency between the two methods.

Comparisons for activity intensity can be found in Figures 3 and 4. In Figure 3, the Sojourn method overestimated time spent sedentary by 26 minutes when compared to the SWA. The Sojourns also overestimated time spent in light intensity activity by 79 minutes, on average. Conversely, the Sojourns consistently underestimated time spent in moderate intensity activity and MVPA by 68 minutes and 54 minutes, respectively. As seen in Figure 4, the Sojourns underestimated time spent sedentary by 113 minutes when compared to the ACT24. A similar underestimation was found by the Sojourns when compared to the ACT24 in moderate intensity (33 minutes) and MVPA (22 minutes). Conversely, the Sojourns overestimated time spent in light intensity activity by 134 minutes. These findings were reflected when observing limits of agreement; the window for MVPA was much narrower than the window for light intensity physical activity, indicating greater consistency between the methods for MVPA.

CHAPTER 5. DISCUSSION

This study aimed to compare multiple methods of capturing time spent at various intensities of activity and energy expenditure in able-bodied adults. In recent years, there has been a noticeable trend towards machine-learning techniques for capturing energy expenditure compared to other methods (Troiano, McClain, Brychta, & Chen, 2014). Thus, a primary goal of the current study was to evaluate how the estimates from the Sojourn method compare with two other alternative indicators. The Sensewear Armband is a commonly used objective monitoring tool and the ACT24 provides a comparison with a report-based measure. The agreement for estimates of energy expenditure are summarized first, followed by exploration of agreement for different intensities of physical activity.

Overall, the estimates of EE from the Sojourn method were quite similar to that of the Sensewear. The MPE of 3.5% reveals a minimal degree of underestimation and the MAPE value of 16.1% shows that estimates are within 16% of each other (on average). The measures were also moderately correlated ($r = 0.65$) with each other. The correlations were stronger between the Sojourn method and the ACT24 ($r = 0.91$) and similarly small indicators of MPE (6.6% Sojourn method underestimation) and MAPE (9.5%). The fact that the Sojourn yielded comparable results to these two different indicators is noteworthy. While neither comparison can be considered to be a true criterion, it is useful to know that the estimates from the Sojourn are similar to the values from these other instruments.

Previous research has supported the utility of machine-learning methods (such as the Sojourn method) specifically for capturing energy expenditure. A recent review of machine-learning methods by Liu, Gao, & Freedson (2012) found that generally artificial neural networks performed comparatively, if not more favorably, compared to various accelerometers for capturing energy expenditure. Similarly, Ellingson, Schwabacher, Kim, Welk, and Cook (2016) found that the Sojourn method performed comparably to their Sojourn-Including-Posture (SIP) method as well as an Oxycon Mobile for capturing energy expenditure. Specifically, the Sojourn method demonstrated an overall MAPE value of 21.9% for capturing energy expenditure, compared to 17.7% for the SIP method. These findings indicate that though the Sojourn method performs comparably to established methods, it can be improved to be a more accurate and precise measure of caloric expenditure.

Though the Sojourn method performed reasonably well for capturing caloric expenditure, its performance varied by intensity when comparing the methods by time allocation. Specifically, the Sojourn method performed comparably to the other methods when assessing time spent sedentary, in vigorous intensity, and in MVPA. Moderate strength correlations were found when comparing the Sojourn method to the other two methods, ranging from .36 to .69. In addition, most of the error values were below 30%; sedentary time error values were all below 17.1%, while error values for MVPA ranged from 17.6% to 77.2%. These findings share similarities with previous research, including the original Lyden et al. (2014) study on the development of the Sojourn method. In that study, where participants were observed for ten consecutive hours in a free-living environment three different times, percent bias for sedentary time was 8.2%, while percent bias for MVPA was a bit higher 72.8%. Such findings indicate that an advantage of the Sojourn method may be capturing activity (or inactivity) at these intensities. Of particular importance is its ability to capture sedentary time and MVPA since these indicators have each been found to have relevance for public health research. The interpretation of the amount of time in the light activity category is still challenging. Shifts from sedentary to light can be beneficial but it can also reflect time that is not in MVPA.

There is currently considerable interest in teasing out the health relevance of time allocation and this has led to interest in concepts referred to as time-use epidemiology. An emerging research group known as the International Network of Time Use Epidemiologists (http://www.intue.org/) has even been established to systematically understand how allocations in one category influence interpretations of other categories. A specific analytic technique known as isotemporal substitution has been used to directly study the health effects associated with substituting one type of activity for another (Mekary et al., 2009). This model has been utilized in numerous studies. For example, Mekary et al. (2013) observed changes in depression risk with various activity displacements, specifically television watching. Replacing 60 minutes of television watching (relative risk = 1.18, 95% confidence interval: 1.05, 1.31) with 60 minutes a day of slow walking was found not be associated with depression risk. However, replacing 60 minutes television watching with 60 minutes of brisk walking was associated with lower depression risk (relative risk = 0.85, 95% confidence interval 0.76, 0.95). Similarly, a study by Lerma et al. (2018) examined the effects of substituting sedentary behavior with light physical activity and MVPA on various measures of physical function, including a 400-m walk test, gait speed, and the five times sit-to-stand test. Replacing 30 minutes of sedentary behavior with light physical activity was associated with significant improvement in the 400-m walk test, while replacing with MVPA was associated with significant improvement in all 3 tests. Thus, the allocation of time in one category of activity may influence a wide variety of outcomes, including agreement within different categories of activity.

The present study does not directly contribute to time use epidemiology research, but it does help to understand the variability in estimates obtained from different instruments over a 24-hour period. The relative accuracy of the assessments is important to identify physical activity trends, to recognize how much physical activity is enough to achieve health benefits, or how sedentary one can be before experiencing its detrimental effects. Thus, there is a need for continuous testing and improvement of methods that capture physical activity to inform such research. This study suggests that the Sojourn method may perform well in that capacity. More advanced iterations of the Sojourn method (e.g. SIP) may further enhance accuracy and precision, but the current version performs comparably to other methods for capturing the same indicators.

Despite its promise for capturing sedentary time, vigorous intensity activity, and MVPA, the Sojourn method showed greater error relative to the SWA and the ACT24 for estimates of moderate and light intensity activity. For moderate intensity activity, the Sojourn method showed a moderate-strength relationship with the SWA ($r = .60$) but little relationship with the ACT24 ($r = .17$). Error values reflect this finding; MPE and MAPE values were vastly different when comparing the Sojourns and SWA (MPE: 37.5%, MAPE: 59.2%) to the Sojourns and ACT24 (MPE: 83.4%, MAPE: 134.8%). For light intensity activity, this trend switches. The Sojourn method shows a moderate strength relationship with the ACT24 ($r = .38$) but a weak correlation with the SWA ($r = .05$) for activity at this intensity. Similarly, error values are higher for the Sojourns and SWA (MPE: 63.6%, MAPE: 82.8%) than they are for the Sojourns and ACT24 (MPE: 33.9%, MAPE: 50.0%). The original results from Lyden et al. (2014) reported a small 8.2% bias for estimation of light intensity activity but this could be due to the nature of the activities used in the more controlled evaluation and the use of a true criterion. It is also important to emphasize that misclassification in one activity will automatically influence classification in another. It is likely that there was considerable overlap in the allocations of light and moderate intensity and this could explain the poor agreement. It is also possible that the nature of the sample could influence the results. The Lyden, et al. (2014) study used a small (n = 13), highly active sample while the current study used a larger sample (n = 85), largely split between young, high-active participants and older, low-active participants. As evidenced by clustered bar graphs seen in Figure 1, there was a considerable amount of light intensity along with MVPA. The range of activity levels and the free-living evaluation in the present study provides a more robust evaluation.

Another compelling finding resulting from this study relates to the use of MPE and MAPE as statistical proxies of error. The primary focus of these analyses was to compare the Sojourn method and the SWA, but additional comparisons were drawn between the SWA and ACT24. As seen in Table 7, MPE values between the two methods were 2.1% and 6.7% for moderate intensity activity alone and MVPA, respectively. At surface level, this would indicate acceptable agreement between the two methods. However, the computed MAPE values reveal the true magnitude of the error at the individual level with MAPE values of 93.9% for moderate intensity activity and 84.8% for MVPA. It is not clear why methods overestimate for some people for other people but the disparity between MAPE and MPE clearly indicates that this is the case.

The difference between MPE and MAPE can be understood can be best understood by understanding the differences between the related terms of “accuracy” and “precision”. While often used interchangeably, it is possible for a measure to be precise but still inaccurate. It is also possible for a measure to be accurate but imprecise. In the present analyses, the correct interpretation is that the Sojourn method is not precise at the individual level, but the overall estimate is reasonably accurate at the group level. In some cases, an accurate group level estimate is all that is needed; however, in some lines of research it is important to be able to estimate individual levels of PA and EE with more precision. Thus, researchers should take care when interpreting the relative importance of MPE and MAPE within a study.

A useful way of visualizing this error is through Bland-Altman plots. This study made use of such plots, which can be seen in Figures 2-4. One point of interest in these plots is the line that represents the mean difference between the two methods. This line allows for the determination of bias of the comparison method (the Sojourns) in reference to the established method (SWA). If this line was close to or near zero, the group difference (or error) between the two methods would be minimal. As seen in Figure 2, which compares the methods on energy expenditure, each is generally within about 200 kilocalories of the others. Thus, if a researcher is willing to accept this magnitude of a difference when studying energy expenditure, then the methods would be considered comparable. However, in Figure 3, which detailed activity at different intensities for the Sojourn method and SWA, differences ranged from 26-79 minutes. Likewise, Figure 4 showed differences between the Sojourn method and ACT24, ranging from 21-134 minutes. These differences are considerable in the field of activity assessment research; it could mean the difference between meeting recommendations or not meeting them. Thus, researchers may use caution when using these methods to assess physical activity behavior, depending on their specific indicator of choice.

Another relevant point of interest in these plots is the slope of the line along with the spread of the data points. As seen in Figure 2, each slope is relatively flat and data points are spread out, indicating that variability is mostly random. However, as seen in the moderate intensity and MVPA plots in Figure 3, there is a clear positive slope to these residuals, suggesting a defined relationship. In both cases, as the amount of time spent at those intensities increase, the magnitude of difference between the Sojourn method and SWA gets larger. Thus, there is more variability at higher intensities than at lower ones. Similar patterns were evident in the parallel plots with Act24 shown in Figure 4. The consistency of these relationships supports the generalizability of the findings, but it is important to note that these relationships could also be influenced by some outlier points. For example, most participants engaged in some moderate intensity activity or MVPA, but some engaged in higher amounts of activity at these intensities (200 minutes and above). Fewer data points at these intensities may have led to less stable findings and they could have also lead to the observed bias in the Bland-Altman plots. More research is needed to fully understand the agreement between these methods, as well as possible sources of error.

Summary, Limitations, and Conclusions

This study aimed to examine the performance of the Sojourn method compared to the Sensewear Armband and ACT24 on various indicators. While the Sojourn method has received considerable attention in the literature there are few other studies that have explored the accuracy of the estimates against other indicators. This study filled this gap by evaluating the performance of the machine-learning Sojourn method for processing triaxial hip Actigraph data against an established monitor-based approach in the Sensewear Armband (SWA) as well as a report-based measure in the ACT24. The study utilized a free-living design, which allowed participants to wear a hip-placed Actigraph as well as a Sensewear Armband over a period of 24 hours, followed by a lab-administered ACT24 evaluation. While the 24-hour period is limited, it provides a systematic way to examine agreement among competing methods under real-world conditions. Another strength of the study was the comprehensive evaluation of agreement that included multiple indicators, including correlations, mean percent error (MPE), mean absolute percent error (MAPE), and Bland-Altman plots. This type of design and evaluation has numerous advantages and advances the field of adult physical activity assessment. The generally good agreement for the Sojourn method for estimates of EE support the continued use of this approach for future research. Results revealed high correlations and reasonably low values for MPE and MAPE against both the SWA and ACT24 so that supports the overall validity of the methodology. However, the lower levels of agreement for categorizing intensity of physical activity suggest the need for further refinement. As pointed out above, there are inherent challenges in this regard since error in one category will automatically introduce error in another category. It is important to consider this when interpreting the results. Overall, the study design allowed for exploration of the advantages (and drawbacks) of using the Sojourn method for processing physical activity data as well as the accuracy in which it captures energy expenditure and time spent at various intensities of activity. While results are specific for the Sojourn method, the approach used to examine the agreement provides a good model for future work with different indicators.

Despite the strengths, the study also has several limitations. The sample itself is a young and active one; 80% of the sample were aged between 18-30. This age group is generally more active compared to their older peers. Thus, they may engage in more MVPA, in which the Sojourn method performed reasonably. An older sample may have engaged in less MVPA, and with Sojourn method showing considerable error with light intensity activity, overall interpretation of the Sojourn method may be different. As a result, though findings are generalizable to that specific age and activity group, caution should be used when generalizing these findings to the entire able-bodied population. In addition, the number of cases dropped from the final analysis are a concern. Though drop out was low and wear time was high, unforeseen complications at random with monitors as well as the self-report process led to missing data for some segments.

It is important to note that the agreement reported here are relative to values from measures that cannot be considered ‘criterion values’. The SWA has an established reputation as perhaps one of the most valid indicator of energy expenditure with multiple studies documenting MAPE values of less than 10% for comparisons with doubly labelled water (Johannsen et al., 2010; Calabro, Kim, Franke, Stewart, & Welk, 2015). However, it is important to note that these comparisons are for total EE which is primarily attributable to sedentary and light activity accumulated throughout the day. Several studies have documented that the estimates of PA from the SWA tend to be higher than values from other methods, but this may be attributable to the internal calculations used to segment the data into different intensities. A recent review of the monitor by Koehler and Drenowatz (2017) noted that the Sensewear tends to overestimate energy expenditure at lower intensities of activity and underestimate energy expenditure at higher intensities of activity. This error would automatically influence agreement between the methods for capturing energy expenditure and should be considered when interpreting the results of this study.

While the focus of the study was on the Sojourn method, it also provides some perspective on the differences between estimates from monitor-based and report-based methods. Typically, report-based methods are limited by questionable reliability and validity, so they are often supplemented with objective measures to reduce the impact of this bias (Prince et al., 2008). Within the physical activity assessment field, the agreement between self-report and objective measures vary. A systematic review by Prince et al. (2008) demonstrated no detectable pattern within differences; self-report of physical activity was greater than directly observed physical activity in some cases, and lower in others. This could potentially influence agreement and lead to error in the current study when comparing the Sojourn method and SWA with the ACT24. However, the previous version of the ACT24, in the form of the 24PAR, was found to be strongly correlated with the Sensewear for energy expenditure and moderately correlated for MVPA (Calabro et al., 2009). Similar results were also reported with an interviewer administered version of the 24PAR (Welk et al., 2014). The similar associations between the online version of the Act24 and the SWA in the present study shows that this 24-hour recall format has strong utility. However, a limitation of the current ACT24 system is that it was not possible to extract the temporally coded activities. This limited the ability to use the ACT24 data to explore factors that may contribute to sources of error.

Another major limitation with the present analyses was the inability to fully separate out sleep time from overall sedentary time. Thus, more time was allocated into the sedentary category for all indicators. There is considerable research underway to better isolate sleep from sedentary, but this was beyond the scope of the present study. Less variation between the methods may have been seen during sleep time, which would have been reflected in overall sedentary time. Had they been separated, findings related to sedentary time may have been different. Thus, future research should look to split the two in order to get a more accurate picture of sedentary time as well as assess the performance of the Sojourn method for capturing sleep.

Despite these limitations, the study offers many insights. The Sojourn method performed well for capturing energy expenditure in relation to the SWA and ACT24. In addition, it performed reasonably for capturing time spent at different intensities of activity. Future research should look to build upon and improve the Sojourn method for capturing light and moderate intensity activity as well as MVPA. Finally, this study exemplifies the difference in interpretation that can be drawn from data using different methods of analysis. Researchers should take care in the statistical methods they use and make sure they are appropriate for the question(s) they are looking to answer.

TABLES AND FIGURES

Table 1. Baseline Demographic and General Health Characteristics

Characteristic	N	%	Mean ± SD
Age	85		26.2 ± 9.9
18-30	68	80	
31-49	14	16.5	
50-60	3	3.5	
Sex	85		
Male	31	36.5	
Female	54	63.5	
Race/ethnicity	85		
White not Hispanic	65	76.5	
White Hispanic	8	9.4	
Black not Hispanic	4	4.7	
Black Hispanic	1	1.2	
Asian	5	5.9	
American Indian (Native American)	0	0	
Other	2	2.3	
Employment Status	85		
Full time	24	28.2	
Part time	36	42.4	
Unemployed, looking for work	11	12.9	
Unemployed, due to health	0	0	
Retired (for any reason)	0	0	
Never worked outside the home	0	0	
Other	14	16.5	
Marital status	85		
Married	16	18.8	
Divorced	1	1.2	
Never married	67	78.8	
Widowed	0	0	
Separated	0	0	
Living as married	1	1.2	
Income	85		
Less than $25,000	50	58.8	
$25,000-$49,999	10	11.8	
$50,000-$74,999	7	8.2	
$75,000-$99,999	9	10.6	
$100,000 or more	9	10.6	
Education	85		
Did not finish High School	0	0	
High School Graduate (or with GED)	11	12.9	
Some College (post-high school AA, vocational degree)	31	36.5	

Table 1. (continued)

College Graduate (Bachelors, 4-year degree)	12	14.1	
Some Postgrad	6	7.1	
Graduate Degree (MS, MA, PhD, MD, etc.)	25	29.4	
Previous use of activity monitor	85		
Yes, currently using	24	28.2	
Yes, not currently but in the last 6 months	14	16.5	
Yes, more than 6 months ago	18	21.2	
No	29	34.1	
Currently physically active	85		
No	8	9.4	
Yes	77	90.6	
Regular physical activity	85		
No	18	21.2	
Yes	67	78.8	
Height (cm)	85		174 ± 9.4
Weight (kg)	85		76.7 ± 16.9
BMI	85		25.3 ± 5.1
Underweight (less than 18.5)	0	0	
Normal weight (18.5-24.9)	44	51.8	
Overweight (25-29.9)	33	38.8	
Obese (30 or more)	8	9.4	

Table 2. Average Activity Level by Method and Sex

Characteristic	N	Mean ± SD
Sojourns		
EE (kcal)	85	2665.2 ± 745.0
Male	31	3004.9 ± 862.2
Female	54	2470.2 ± 593.8
Sedentary (min)	85	1015.3 ± 145.1
Male	31	987.0 ± 179.1
Female	54	1025.4 ± 128.3
Light PA (min)	85	336.9 ± 140.5
Male	31	358.4 ± 163.4
Female	54	329.5 ± 130.1
Moderate PA (min)	85	53.7 ± 25.1
Male	31	60.4 ± 27.2
Female	54	51.5 ± 26.4
Vigorous PA (min)	85	34.1 ± 29.7
Male	31	34.2 ± 24.7
Female	54	33.6 ± 32.4
MVPA (min)	85	87.9 ± 43.2
Male	31	94.6 ± 37.4
Female	54	85.1 ± 46.9
Sensewear Armband		
EE (kcal)	85	2869.0 ± 788.7
Male	31	3282.5 ± 526.4
Female	54	2526.1 ± 579.8
Sedentary (min)	85	989.1 ± 181.9
Male	31	916.2 ± 194.8
Female	54	1030.9 ± 161.4
Light PA (min)	85	258.0 ± 94.2
Male	31	283.4 ± 84.1
Female	54	243.5 ± 97.4
Moderate PA (min)	85	121.7 ± 70.2
Male	31	136.4 ± 59.5
Female	54	113.2 ± 74.9
Vigorous PA (min)	85	20.0 ± 30.3
Male	31	21.1 ± 28.8

Table 2. (continued)

Female	54	19.4 ± 31.3
MVPA (min)	85	141.6 ± 89.1
Male	31	157.4 ± 70.7
Female	54	132.6 ± 97.6
ACT24		
EE (kcal)	85	2802.0 ± 667.2
Male	31	3245.6 ± 874.4
Female	54	2652.9 ± 650.1
Sedentary (min)	85	1128.0 ± 141.8
Male	31	1117.5 ± 163.0
Female	54	1134.1 ± 142.7
Light PA (min)	85	202.6 ± 141.8
Male	31	214.7 ± 154.6
Female	54	195.7 ± 134.9
Moderate PA (min)	85	87.0 ± 102.9
Male	31	76.3. ± 82.0
Female	54	93.2 ± 113.5
Vigorous PA (min)	85	22.4 ± 37.2
Male	31	31.6 ± 46.5
Female	54	17.1 ± 29.8
MVPA (min)	85	109.4 ± 104.5
Male	31	107.9 ± 82.9
Female	54	110.3 ± 115.8

Table 3. Time Spent at Various Intensities: Sojourns and SWA

Method	SWA Sedentary	SWA Light	SWA Moderate	SWA Vigorous	SWA MVPA
Sojourns Sedentary	.38**	-.15	-.16	-.17	-.18
Sojourns Light	-.23*	.05	-.04	.03	-.02
Sojourns Moderate	-.45**	.26*	.60**	.23*	.55**
Sojourns Vigorous	-.41**	.27*	.44**	.46**	.50**
Sojourns MVPA	-.55**	.33**	.65**	.44*	.45**

*Significant at .05 level **Significant at .01 level

Table 4. Time Spent at Various Intensities: SWA and ACT24

Method	SWA Sedentary	SWA Light	SWA Moderate	SWA Vigorous	SWA MVPA
ACT24 Sedentary	.31**	-.22*	-.22*	-.21	-.24*
ACT24 Light	-.13	.12	.03	.10	.06
ACT24 Moderate	-.12	.08	.11	-.09	.05
ACT24 Vigorous	-.42**	.17	.44**	.69**	.58**
ACT24 MVPA	-.27**	.14	.27*	.15	.26*

*Significant at .05 level **Significant at .01 level

Table 5. Time Spent at Various Intensities: Sojourns and ACT24

Method	Sojourns Sedentary	Sojourns Light	Sojourns Moderate	Sojourns Vigorous	Sojourns MVPA
ACT24 Sedentary	.52**	-.46**	-.29**	-.09	-.23*
ACT24 Light	-.36**	.38**	.12	-.14	-.02
ACT24 Moderate	-.21	.15	.17	.16	.21
ACT24 Vigorous	-.12	-.01	.23*	.46**	.45**
ACT24 MVPA	-.25*	.15	.25*	.32**	.36**

*Significant at .05 level **Significant at .01 level

Table 6. EE Correlation

Method	Sojourns	SWA	ACT24
Sojourns	-	.65**	.91**
SWA	-	-	.59**
ACT24	-	-	-

*Significant at .05 level **Significant at .01 level

Table 7. MPE and MAPE Analyses

Indicator and Method	**MPE**	**MAPE**
Energy Expenditure		
Sojourns, SWA	3.5%	16.1%
Sojourns, ACT24	6.6%	9.5%
SWA, ACT24	4.4%	17.1%
Sedentary time		
Sojourns**, SWA**	5.6%	17.1%
Sojourns, ACT24	12.5%	15.6%
SWA, ACT24	18.4%	23.7%
Light intensity		
Sojourns**, SWA**	63.6%	82.8%
Sojourns**, ACT24**	33.9%	50.0%
SWA**, ACT24**	9.6%	55.1%
Moderate intensity		
Sojourns, SWA	37.5%	59.2%
Sojourns, ACT24	83.4%	134.8%
SWA, ACT24	2.1%	93.9%
MVPA		
Sojourns, SWA	17.6%	45.0%
Sojourns, ACT24	30.0%	77.2%
SWA, ACT24	6.7%	84.8%

***Bold** indicates which method was higher

Figure 1. Average Activity Level by Intensity Classification

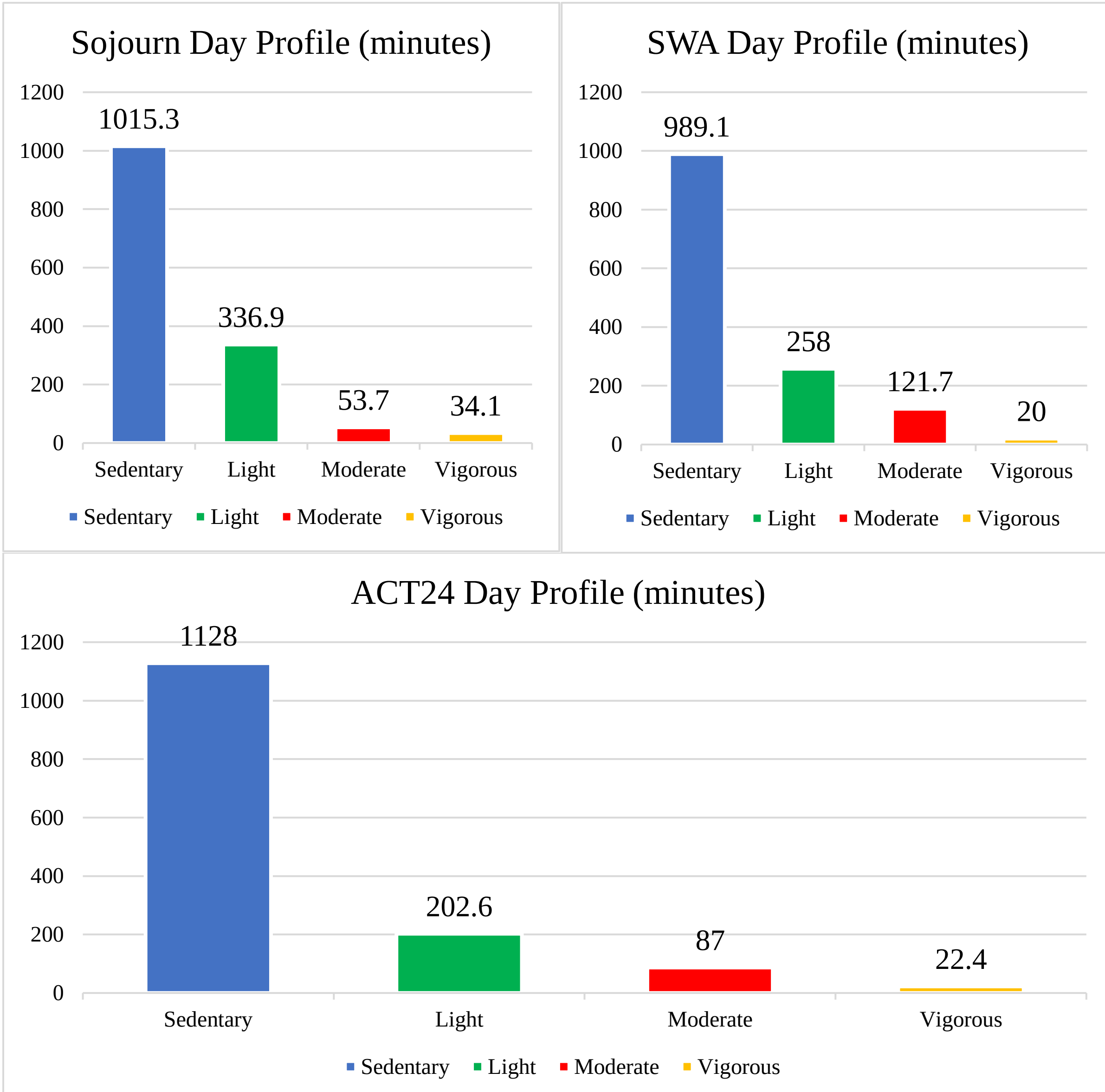


Figure 2. Bland-Altman Plots: Energy Expenditure Estimation

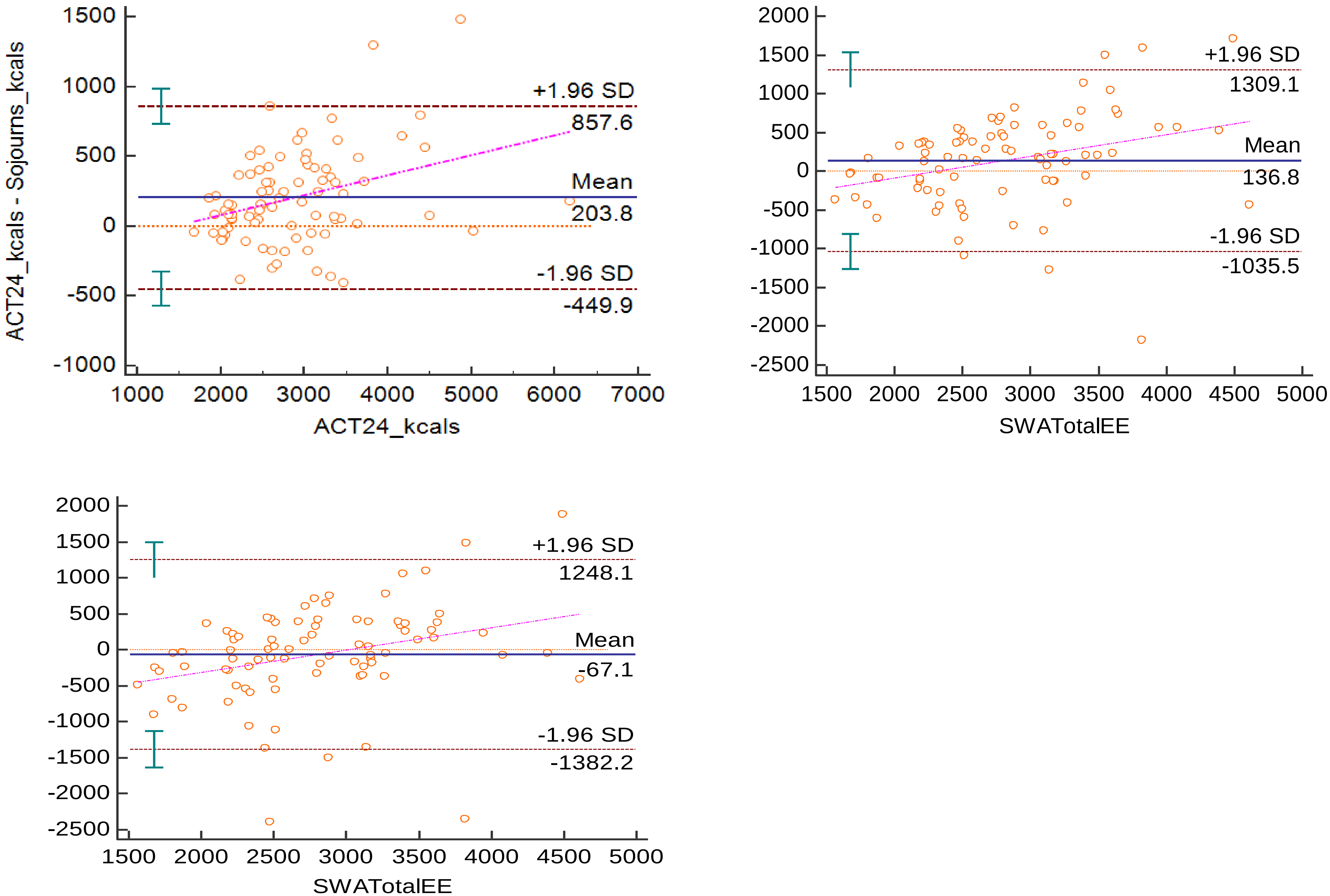


Figure 3. Bland-Altman Plots: Sojourns and SWA

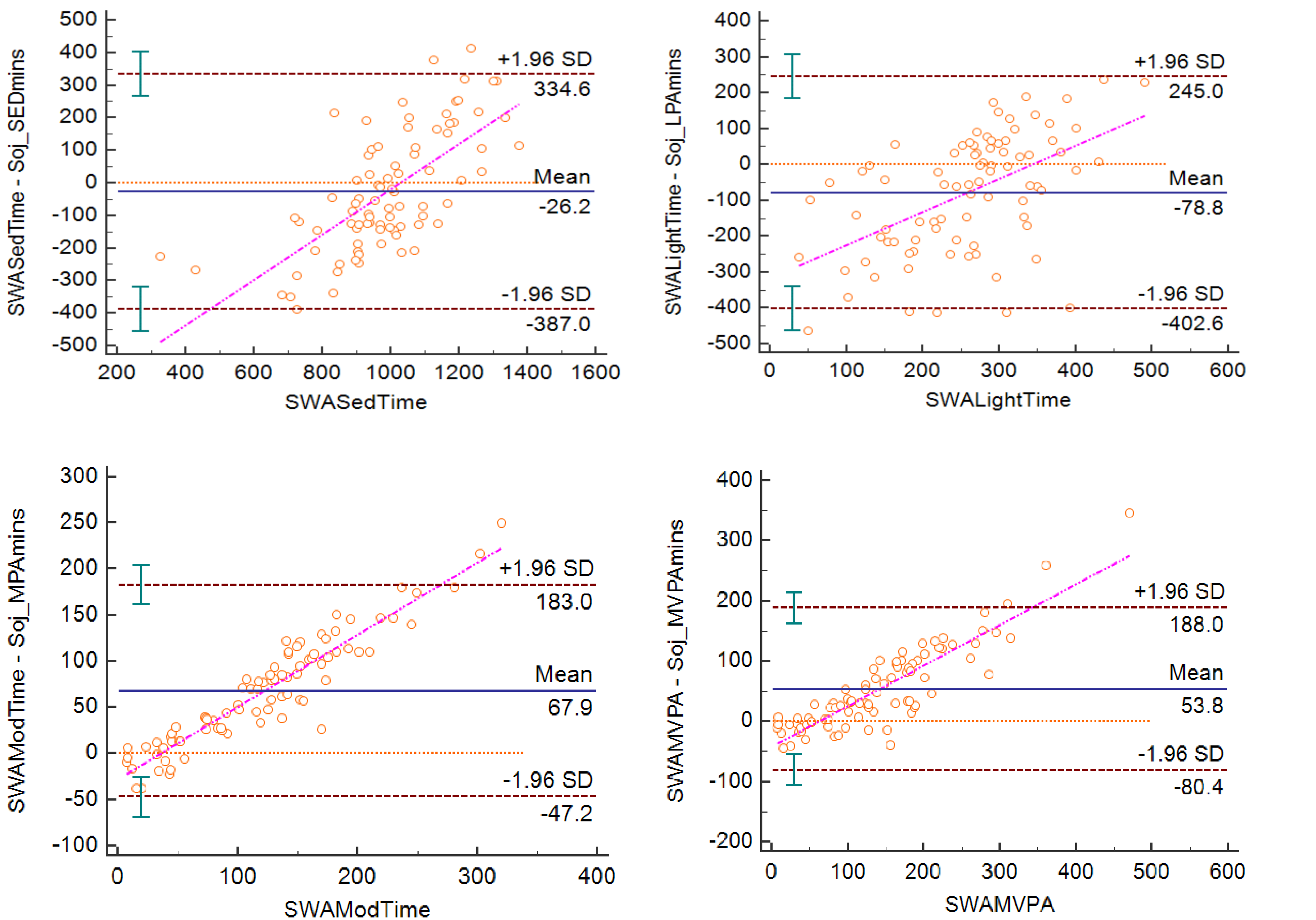
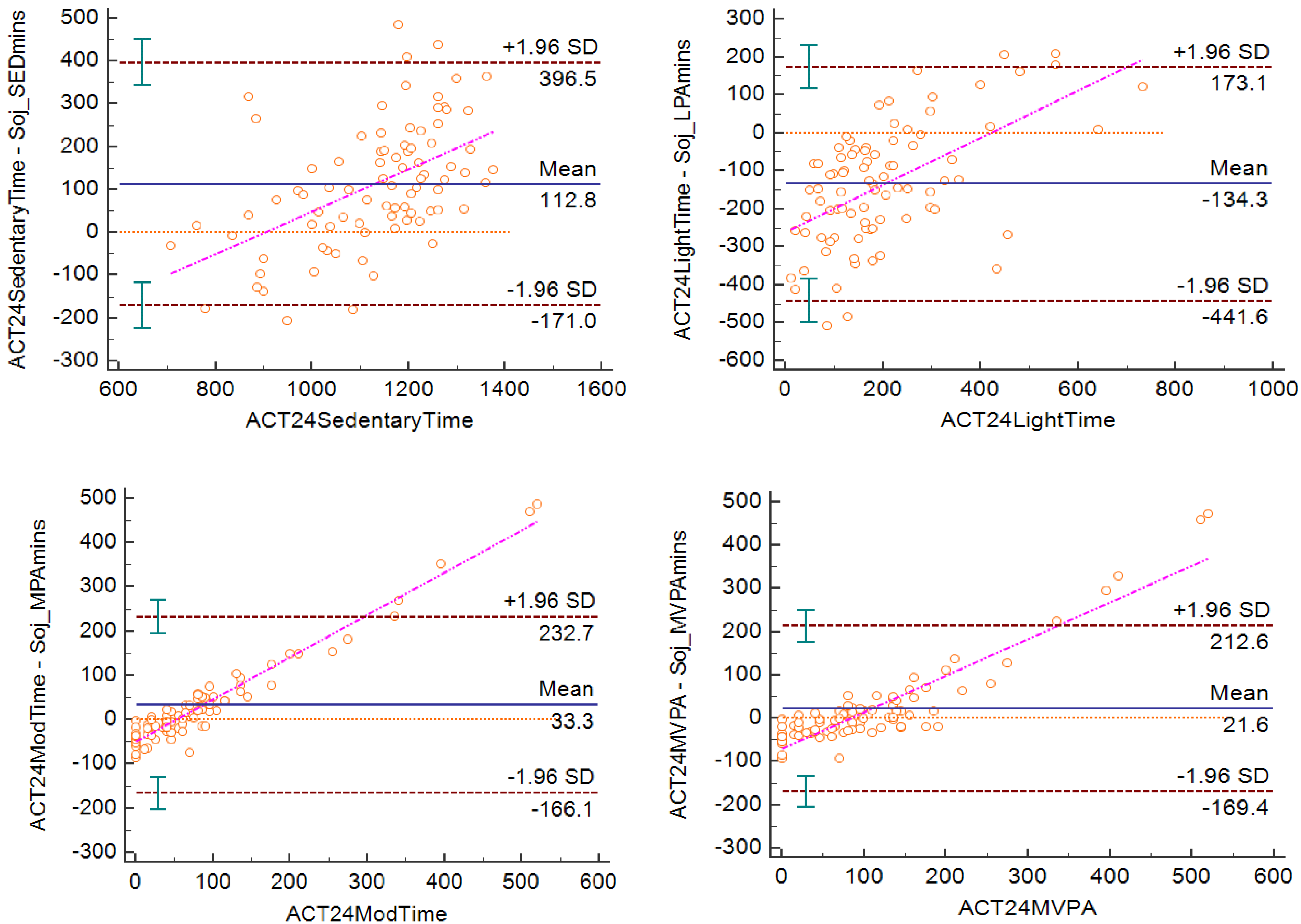


Figure 4. Bland-Altman Plots: Sojourns and ACT24



REFERENCES

Ainsworth, B. E., Haskell, W. L., Herrmann, S. D., Meckes, N., Bassett, D. R., Tudor-Locke, C., Greer, J.L., Vezina, J., Whitt-Glover, M.C., & Leon, A. S. (2011). 2011 Compendium of Physical Activities: A Second Update of Codes and MET Values. *Medicine & Science in Sports & Exercise*, *43*(8), 1575–1581. https://doi.org/10.1249/MSS.0b013e31821ece12

Atkin, A. J., Gorely, T., Clemes, S. A., Yates, T., Edwardson, C., Brage, S., Salmon, J., Marshall, S.J., & Biddle, S. J. (2012). Methods of Measurement in epidemiology: Sedentary Behaviour. *International Journal of Epidemiology*, *41*(5), 1460–1471. https://doi.org/10.1093/ije/dys118

Bassett, D. R., Rowlands, A. V., & Trost, S. G. (2012). Calibration and Validation of Wearable Monitors. *Medicine and Science in Sports and Exercise*, *44*(1 Suppl 1), S32–S38. https://doi.org/10.1249/MSS.0b013e3182399cf7

Blair, S. N., Haskell, W. L., Ho, P., Paffenbarger, R. S., Vranizan, K. M., Farquhar, J. W., & Wood, P. D. (1985). Assessment of habitual physical activity by a seven-day recall in a community survey and controlled experiments. *American Journal of Epidemiology*, *122*(5), 794–804.

Bouchard, C., Tremblay, A., Leblanc, C., Lortie, G., Savard, R., & Thériault, G. (1983). A method to assess energy expenditure in children and adults. *The American Journal of Clinical Nutrition*, *37*(3), 461–467.

Calabro, Miguel A., Welk, G. J., Carriquiry, A. L., Nusser, S. M., Beyler, N. K., & Mathews, C. E. (2009). Validation of a computerized 24-hour physical activity recall (24PAR) instrument with pattern-recognition activity monitors. *Journal of Physical Activity & Health*, *6*(2), 211–220.

Calabro, Miguel Andres, Kim, Y., Franke, W. D., Stewart, J. M., & Welk, G. J. (2015). Objective and subjective measurement of energy expenditure in older adults: A doubly-labeled water study. *European Journal of Clinical Nutrition*, *69*(7), 850–855. https://doi.org/10.1038/ejcn.2014.241

Castillo-Retamal, M., & Hinckson, E. A. (2011). Measuring physical activity and sedentary behaviour at work: a review. *Work (Reading, Mass.)*, *40*(4), 345–357. https://doi.org/10.3233/WOR-2011-1246

CDC. (2015). Physical Activity and Health. Retrieved November 28, 2017, from https://www.cdc.gov/physicalactivity/basics/pa-health/index.htm

Clarke, T. C., Norris, T., & Schiller, J. S. (2016). Early release of selected estimates based on data from the 2016 National Health Interview Survey. *National Center for Health Statistics. Available at Http://www.cdc.gov/Nchs/about/Major/Nhis/About200506.htm*.

Corder, K., Brage, S., & Ekelund, U. (2007). Accelerometers and pedometers: methodology and clinical application. *Current Opinion in Clinical Nutrition and Metabolic Care*, *10*(5), 597–603. https://doi.org/10.1097/MCO.0b013e328285d883

Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., Pratt, M., Ekelund, U., Yngve, A., Sallis, J.F., & Oja, P. (2003). International Physical Activity

Questionnaire: 12-country Reliability and Validity. *Medicine & Science in Sports & Exercise*, *35*(8), 1381–1395. https://doi.org/10.1249/01.MSS.0000078924.61453.FB

Crouter, S. E., Clowers, K. G., & Bassett, D. R. (2006). A novel method for using accelerometer data to predict energy expenditure. *Journal of Applied Physiology*, *100*(4), 1324–1331. https://doi.org/10.1152/japplphysiol.00818.2005

Crouter, S. E., Kuffel, E., Haas, J. D., Frongillo, E. A., & Bassett, D. R. (2010). Refined two-regression model for the ActiGraph accelerometer. *Medicine and Science in Sports and Exercise*, *42*(5), 1029–1037. https://doi.org/10.1249/MSS.0b013e3181c37458

Ellingson, L. D., Schwabacher, I. J., Kim, Y., Welk, G. J., & Cook, D. B. (2016). Validity of an Integrative Method for Processing Physical Activity Data. *Medicine & Science in Sports & Exercise*, *48*(8), 1629. https://doi.org/10.1249/MSS.0000000000000915

Evenson, K. R., Catellier, D. J., Gill, K., Ondrak, K. S., & McMurray, R. G. (2008). Calibration of two objective measures of physical activity for children. *Journal of Sports Sciences*, *26*(14), 1557–1565. https://doi.org/10.1080/02640410802334196

Freedson, P., Bowles, H. R., Troiano, R., & Haskell, W. (2012). Assessment of Physical Activity Using Wearable Monitors: Recommendations for Monitor Calibration and Use in the Field. *Medicine and Science in Sports and Exercise*, *44*(1 Suppl 1), S1–S4. https://doi.org/10.1249/MSS.0b013e3182399b7e

Freedson, P., Pober, D., & Janz, K. F. (2005). Calibration of accelerometer output for children. *Medicine and Science in Sports and Exercise*, *37*(11 Suppl), S523-530.

Freedson, P. S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and Science in Sports and Exercise*, *30*(5), 777–781.

Gabriel, K. P., McClain, J. J., Schmid, K. K., Storti, K. L., High, R. R., Underwood, D. A., Kuller, L.H., & Kriska, A. M. (2010). Issues in accelerometer methodology: the role of epoch length on estimates of physical activity and relationships with health outcomes in overweight, post-menopausal women. *International Journal of Behavioral Nutrition and Physical Activity*, *7*, 53. https://doi.org/10.1186/1479-5868-7-53

Giavarina, D. (2015). Understanding Bland Altman analysis. *Biochemia Medica*, *25*(2), 141–151. https://doi.org/10.11613/BM.2015.015

Harvey, J. A., Chastin, S. F. M., & Skelton, D. A. (2015). How Sedentary Are Older People? A Systematic Review of the Amount of Sedentary Behavior. *Journal of Aging and Physical Activity*, *23*(3), 471–487. https://doi.org/10.1123/japa.2014-0164

Helmerhorst, H. H. J., Brage, S., Warren, J., Besson, H., & Ekelund, U. (2012). A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. *International Journal of Behavioral Nutrition and Physical Activity*, *9*, 103. https://doi.org/10.1186/1479-5868-9-103

Janz, K. F. (1994). Validation of the CSA accelerometer for assessing children’s physical activity. *Medicine and Science in Sports and Exercise*, *26*(3), 369–375.

Johannsen, D. L., Calabro, M. A., Stewart, J., Franke, W., Rood, J. C., & Welk, G. J. (2010). Accuracy of armband monitors for measuring daily energy expenditure in healthy adults. *Medicine and Science in Sports and Exercise*, *42*(11), 2134–2140. https://doi.org/10.1249/MSS.0b013e3181e0b3ff

Kelly, L. A., McMillan, D. G., Anderson, A., Fippinger, M., Fillerup, G., & Rider, J. (2013). Validity of actigraphs uniaxial and triaxial accelerometers for assessment of physical activity in adults in laboratory conditions. *BMC Medical Physics*, *13*, 5. https://doi.org/10.1186/1756-6649-13-5

Kim, Youngdeok, Barry, V. W., & Kang, M. (2015). Validation of the ActiGraph GT3X and activPAL Accelerometers for the Assessment of Sedentary Behavior. *Measurement in Physical Education and Exercise Science*, *19*(3), 125–137. https://doi.org/10.1080/1091367X.2015.1054390

Kim, Youngwon, & Welk, G. J. (2017). The accuracy of the 24-h activity recall method for assessing sedentary behaviour: the physical activity measurement survey (PAMS) project. *Journal of Sports Sciences*, *35*(3), 255–261. https://doi.org/10.1080/02640414.2016.1161218

Koehler, K., & Drenowatz, C. (2017). Monitoring Energy Expenditure Using a Multi-Sensor Device—Applications and Limitations of the SenseWear Armband in Athletic Populations. *Frontiers in Physiology*, *8*. https://doi.org/10.3389/fphys.2017.00983

Kriska, A. M., Knowler, W. C., LaPorte, R. E., Drash, A. L., Wing, R. R., Blair, S. N., Bennett, P.H., & Kuller, L. H. (1990). Development of Questionnaire to Examine Relationship of Physical Activity and Diabetes in Pima Indians. *Diabetes Care*, *13*(4), 401–411. https://doi.org/10.2337/diacare.13.4.401

Lee, J.-M., Kim, Y., Bai, Y., Gaesser, G. A., & Welk, G. J. (2016). Validation of the SenseWear mini armband in children during semi-structure activity settings. *Journal of Science and Medicine in Sport*, *19*(1), 41–45. https://doi.org/10.1016/j.jsams.2014.10.004

Lerma, N. L., Cho, C. C., Swartz, A. M., Miller, N. E., Keenan, K. G., & Strath, S. J. (2018). Isotemporal Substitution of Sedentary Behavior and Physical Activity on Function. *Medicine and Science in Sports and Exercise*, *50*(4), 792–800. https://doi.org/10.1249/MSS.0000000000001491

Liu, S., Gao, R., & Freedson, P. (2012). Computational Methods for Estimating Energy Expenditure in Human Physical Activities. *Medicine and Science in Sports and Exercise*, *44*(11), 2138–2146. https://doi.org/10.1249/MSS.0b013e31825e825a

Lopez, G. A., Brønd, J. C., Andersen, L. B., Dencker, M., & Arvidsson, D. (2017). Validation of SenseWear Armband in children, adolescents, and adults. *Scandinavian Journal of Medicine & Science in Sports*, n/a-n/a. https://doi.org/10.1111/sms.12920

Lyden, K., Kozey-Keadle, S. L., Staudenmayer, J., & Freedson, P. S. (2014). A Method to Estimate Free-Living Active and Sedentary Behavior from an Accelerometer: *Medicine & Science in Sports & Exercise*, *46*(2), 386–397. https://doi.org/10.1249/MSS.0b013e3182a42a2d

MacAuley, D. (1994). A history of physical activity, health and medicine. *Journal of the Royal Society of Medicine*, *87*(1), 32–35.

Mackey, D. C., Manini, T. M., Schoeller, D. A., Koster, A., Glynn, N. W., Goodpaster, B. H., Satterfield, S., Newman, A.B., Harris, T.B., & Cummings, S. R. (2011). Validation of an armband to measure daily energy expenditure in older adults. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *66*(10), 1108–1113. https://doi.org/10.1093/gerona/glr101

Matthews, C. E., Chen, K. Y., Freedson, P. S., Buchowski, M. S., Beech, B. M., Pate, R. R., & Troiano, R. P. (2008). Amount of Time Spent in Sedentary Behaviors in the United States, 2003-2004. *American Journal of Epidemiology*, *167*(7), 875–881. https://doi.org/10.1093/aje/kwm390

Matthews, Charles E., Kozey-Keadle, S., Moore, S. C., Schoeller, D. S., Carroll, R. J., Troiano, R. P., & Sampson, J. N. (2017). Measurement of Active and Sedentary Behavior in Context of Large Epidemiologic Studies. *Medicine and Science in Sports and Exercise*. https://doi.org/10.1249/MSS.0000000000001428

Matthews, Charles E., Moore, S., C., George, S. M., Sampson, J., & Bowles, H. R. (2012). Improving Self-reports of Active and Sedentary Behaviors in Large Epidemiologic Studies. *Exercise and Sport Sciences Reviews*, *40*(3), 118–126. https://doi.org/10.1097/JES.0b013e31825b34a0

Mekary, R. A., Lucas, M., Pan, A., Okereke, O. I., Willett, W. C., Hu, F. B., & Ding, E. L. (2013). Isotemporal Substitution Analysis for Physical Activity, Television Watching, and Risk of Depression. *American Journal of Epidemiology*, *178*(3), 474–483. https://doi.org/10.1093/aje/kws590

Mekary, R. A., Willett, W. C., Hu, F. B., & Ding, E. L. (2009). Isotemporal Substitution Paradigm for Physical Activity Epidemiology and Weight Change. *American Journal of Epidemiology*, *170*(4), 519–527. https://doi.org/10.1093/aje/kwp163

Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Nyström, C. D., Mora-Gonzalez, J., Löf, M., Labayen, I., Ruiz, J.R., & Ortega, F. B. (2017). Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports Medicine*, *47*(9), 1821–1845. https://doi.org/10.1007/s40279-017-0716-0

Miller, N. E., Strath, S. J., Swartz, A. M., & Cashin, S. E. (2010). Estimating Absolute and Relative Physical Activity Intensity Across Age via Accelerometry in Adults. *Journal of Aging and Physical Activity*, *18*(2), 158–170.

Mitchell, N., Catenacci, V., Wyatt, H. R., & Hill, J. O. (2011). Obesity: Overview Of An Epidemic. *The Psychiatric Clinics of North America*, *34*(4), 717–732. https://doi.org/10.1016/j.psc.2011.08.005

Morris, J.N., Heady, J.A., Raffle, P.A.B., Roberts, C.G., & Parks, J.N. (1953). Coronary heart disease and physical activity of work. *Lancet*, 1111-1120.

ODPHP. (2017). Appendix 1. Translating Scientific Evidence About Total Amount and Intensity of Physical Activity Into Guidelines. Retrieved November 15, 2017, from https://health.gov/paguidelines/guidelines/appendix1.aspx

Ogden, C. L., Carroll, M. D., Fryar, C. D., & Flegal, K. M. (2015). *Prevalence of obesity among adults and youth: United States, 2011-2014*. US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.

Owen, N., Healy, G. N., Matthews, C. E., & Dunstan, D. W. (2010). Too Much Sitting: The Population-Health Science of Sedentary Behavior. *Exercise and Sport Sciences Reviews*, *38*(3), 105–113. https://doi.org/10.1097/JES.0b013e3181e373a2

Paffenbarger, R. S., Blair, S. N., & Lee, I.-M. (2001). A history of physical activity, cardiovascular health and longevity: the scientific contributions of Jeremy N Morris, DSc, DPH, FRCP. *International Journal of Epidemiology*, *30*(5), 1184–1192. https://doi.org/10.1093/ije/30.5.1184

Physical Activity Guidelines Advisory Committee. (2018). *Physical Activity Guidelines Advisory Committee Scientific Report*. Washington, DC: U.S. Department of Health and Human Services, 2018.

Prince, Stéphanie A, Adamo, K. B., Hamel, M. E., Hardt, J., Gorber, S. C., & Tremblay, M. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *The International Journal of Behavioral Nutrition and Physical Activity*, *5*, 56. https://doi.org/10.1186/1479-5868-5-56

Prince, Stephanie A., Reed, J. L., Mark, A. E., Blanchard, C. M., Grace, S. L., & Reid, R. D. (2015). A Comparison of Accelerometer Cut-Points among Individuals with Coronary Artery Disease. *PLOS ONE*, *10*(9), e0137759. https://doi.org/10.1371/journal.pone.0137759

Sasaki, J. E., John, D., & Freedson, P. S. (2011). Validation and comparison of ActiGraph activity monitors. *Journal of Science and Medicine in Sport*, *14*(5), 411–416. https://doi.org/10.1016/j.jsams.2011.04.003

Tremblay, M. S., Aubert, S., Barnes, J. D., Saunders, T. J., Carson, V., Latimer-Cheung, A. E., Chastin, S.F., Altenburg, T.M., & Chinapaw, M. J. M. (2017). Sedentary Behavior Research Network (SBRN) – Terminology Consensus Project process and outcome. *International Journal of Behavioral Nutrition and Physical Activity*, *14*, 75. https://doi.org/10.1186/s12966-017-0525-8

Treuth, M. S., Schmitz, K., Catellier, D. J., McMurray, R. G., Murray, D. M., Almeida, M. J., Going, S., Norman, J.E., & Pate, R. (2004). Defining Accelerometer Thresholds for Activity Intensities in Adolescent Girls. *Medicine and Science in Sports and Exercise*, *36*(7), 1259–1266.

Troiano, R. P., McClain, J. J., Brychta, R. J., & Chen, K. Y. (2014). Evolution of accelerometer methods for physical activity research. *Br J Sports Med*, *48*(13), 1019–1023. https://doi.org/10.1136/bjsports-2014-093546

Trost, S. G., Ward, D. S., Moorehead, S. M., Watson, P. D., Riner, W., & Burke, J. R. (1998). Validity of the computer science and applications (CSA) activity monitor in children. *Medicine and Science in Sports and Exercise*, *30*(4), 629–633.

Trost, Stewart G., Loprinzi, P. D., Moore, R., & Pfeiffer, K. A. (2011). Comparison of Accelerometer Cut Points for Predicting Activity Intensity in Youth: *Medicine & Science in Sports & Exercise*, *43*(7), 1360–1368. https://doi.org/10.1249/MSS.0b013e318206476e

Vanhelst, J., Béghin, L., Duhamel, A., Bergman, P., Sjöström, M., & Gottrand, F. (2012). Comparison of uniaxial and triaxial accelerometry in the assessment of physical activity among adolescents under free-living conditions: the HELENA study. *BMC Medical Research Methodology*, *12*, 26. https://doi.org/10.1186/1471-2288-12-26

Watson, K. B., Carlson, S., Carroll, D. D., & Fulton, J. (2014). Comparison of Accelerometer Cut Points to Estimate Physical Activity in U.S. Adults. *Journal of Sports Sciences*, *32*(7), 660–669. https://doi.org/10.1080/02640414.2013.847278

Welk, Gregory J., Kim, Y., Stanfill, B., Osthus, D. A., Calabro, A. M., Nusser, S. M., & Carriquiry, A. (2014). Validity of 24-h Physical Activity Recall: Physical Activity Measurement Survey. *Medicine and Science in Sports and Exercise*, *46*(10), 2014–2024. https://doi.org/10.1249/MSS.0000000000000314

Wickel, E. E., Welk, G. J., & Eisenmann, J. C. (2006). Concurrent validation of the Bouchard Diary with an accelerometry-based monitor. *Medicine and Science in Sports and Exercise*, *38*(2), 373–379. https://doi.org/10.1249/01.mss.0000183344.46077.bb