2012

Methods of physical activity assessment in older adults

Colleen Marie Campbell
Iowa State University

Methods of physical activity assessment in older adults

by

Colleen Campbell

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Kinesiology (Behavioral Basis of Physical Activity)

Program of Study Committee:
Warren D. Franke, Major Professor
Gregory J. Welk
Lorraine Lanningham-Foster
Marian Kohut

Iowa State University

Ames, Iowa

2012

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to thank the participants in this study,

my friends in the Kinesiology Department at ISU for their help,

my POS committee,

my parents for their support and patience,

and my brother Tim, as his constant proof-reading, scientific input, and support were invaluable to this project.

ABSTRACT

Introduction: This study examined the validity of several methods of physical activity assessment in older adults. The physical activity questionnaires (PAQs) used included two designed for use with older adults known as the Yale Physical Activity Survey (YPAS) and the Physical Activity Scale for the Elderly (PASE) and one designed for use in the general population, the 7 Day Physical Activity Recall (PAR). These PAQs along with one accelerometer, the ActiGraph GT1M (GT1M), were examined in comparison to the criterion method, the SenseWear Pro 3 Armband (SP3).

Methods: Participants (n = 36; age = 69.7 ± 5.9 years) wore the SP3 and the GT1M for seven days. At the end of the seven days, participants completed the YPAS, PASE, and PAR. Pearson and Spearman correlation coefficients were used to examine significant relationships. Paired samples t-tests and Bland-Altman plots investigated further differences between methods.

Results: GT1M's estimates of daily steps were not significantly correlated (r = .369) with SP3 assessment of daily steps. The ACT-F (r = .620) and ACT-C (r = .524) equations each demonstrated a significant ($P<0.01$) correlation with SP3 estimates of physical activity energy expenditure (PAEE). Bland-Altman analysis indicated that the ACT-C underestimated PAEE by an average of 184 kcal/day. The PAR was the only one PAQ to demonstrate any significant associations with the SP3. These relationships include the PAR PAEE with SP3 PAEE (r = 4.64, $P\leq0.05$), SP3 total energy expenditure (TEE) (r = .556, $P\leq0.01$), SP3 physical activity (PA) (r = .394, $P\leq0.05$) and the PAR PA with SP3 PA (r = .376, $P\leq0.05$). Bland-Altman plot assessments indicated that the PAR overestimated PAEE (1007 kcal/day) and PA (1294 min/week) more than

either the PASE (PA = 324 min/week) or YPAS (PAEE = 476 kcal/week & PA = 230 min/week). The PAR also demonstrated a very strong systematic bias.

Discussion: While the GT1M did not show a significant correlation between estimates of step count with the SP3, it correlated reasonably with other estimates from the SP3 and provided the closest estimates in comparison to the SP3. Thus, it is a reasonable choice when assessing PA in older adults, but may not always be practical due to its high initial cost. As for the PAQs, while the PAR demonstrated the highest correlations, it also demonstrated high systematic bias and PA estimates were the furthest from average SP3 estimates of all methods examined. While PASE and YPAS provided slightly more accurate estimates, their correlations were not significant in comparison to the SP3. Estimates with PAQs were more accurate in individuals with lower PA levels, however may still be useful when used in those with higher PA levels to determine PA patterns.

CHAPTER 1. INTRODUCTION

Physical activity helps prevent Type II diabetes mellitus (Knowler et al., 2002), osteoporosis (Kohrt et al., 2004), stroke (Wendel-Vos et al., 2004), and coronary heart disease (Oguma & Shinoda-Tagawa, 2004) while physical inactivity has been linked to declines in both muscle strength (Hunter, McCarthy, & Barnman, 2004) and cardiovascular function (American College of Sports Medicine Position Stand, 1998). Yet with all the evidence of the benefits of physical activity, 61% of adults over the age of 65 are not meeting the recommended levels of physical activity (Centers for Disease Control Behavioral Risk Factor Surveillance System, 2007). The number of adults over 65 years of age in the United States has increased from 31.2 million in 1990 to 35 million in 2000 and now accounts for 12.4% of the total population (Hetzel & Smith, 2001). Life expectancy has now increased to nearly 78 years of age (Arias, 2007). Increased physical activity can lead to an increased independence and higher quality of life by helping combat medical conditions often associated with age. Thus, it is vital to target the older adult population with activity promotion. In order to decrease the number of elderly who are inactive, it is important to quantify the types of physical activity they are participating in, including assessing types, intensities, durations, and frequency of the activities. It is also important to be able to assess physical activity accurately and reliably in order to better understand the relationship between physical activity and health (Bonnefoy et al., 2001).

There are many ways to assess physical activity, including self-report questionnaires, physical activity monitors, and the doubly labeled water (DLW) technique. DLW is the most accurate method for assessing average energy expenditure over a period of time and therefore often is considered to be the “gold standard.” However, DLW is only able to assess average energy expenditure over a period of time. It is also unable to provide information on physical activity patterns throughout the day. Due to its high cost, it is impractical for everyday use (Welk, 2002). Physical activity questionnaires are used most frequently for physical activity assessment for many reasons. These include their ease of administration, low cost, and ability to capture quantitative and qualitative data (Welk, 2002). There are disadvantages to questionnaire use, including reliability and validity issues associated with difficulty in reliable activity recall. Moreover, the ability of questionnaires that were developed using younger populations to assess physical activity trends in older populations is uncertain (Washburn, 2000). Physical activity questionnaires are likely more accurate in the recall of higher intensity activities than lower intensity activities (Bonnefoy et al., 2001). Older adults typically engage in light and moderate physical activity. Due to the decreased ability of general questionnaires to assess physical activity accurately in older populations, there have been several questionnaires developed for this population. These include the Physical Activity Scale for the Elderly (PASE) (Washburn, Smith, Jette, & Janney, 1993) and the Yale Physical Activity Survey for the Elderly (YPASE) (Dipietro, Caspersen, Ostfeld, & Nadel, 1993).

As with questionnaires, there are many different types of activity monitors. These include accelerometers such as the BioTrainer (IM Systems), Tritrac (Stayhealthy Inc.), and Actigraph (Actigraph LLC). The Actigraph (GT1M) is the most widely used accelerometer (Welk, 2002) and widely available to the general public. Recently there has been a new type of activity monitor developed, the SenseWear Pro 3 armband (SP3), which not only incorporates accelerometry into its energy expenditure calculations, but also heat flux, galvanic skin response, skin temperature, and near body temperature via a variety of sensors located on the device (Jakicic et al., 2004). It also takes into account gender, age, height and weight in the energy expenditure algorithms (Jakicic et al, 2004). It is a relatively new type of monitor, but has been shown to be highly valid and reliable (Jakicic et al., 2004; Fruin & Rankin, 2004). With these design patterns, the SP3 may provide a more accurate estimate of energy expenditure than more conventional accelerometers (Jakicic, et al., 2004; Fruin & Rankin, 2004). However, there are some problems that have arisen with the SP3, such as its ability to accurately apply the appropriate energy expenditure algorithms under free living conditions (Jakicic et al., 2004).

Due to the uncertainties associated with use of physical activity questionnaires in older adults and their ability to not only accurately assess energy expenditure but also physical activity patterns, the proposed study will evaluate the ability of two questionnaires specifically developed for the older adult population. These include the PASE and the YPAS as well as a questionnaire used extensively in the general population known as the Stanford 7-Day (PAR), to assess energy expenditure in older adults accurately. To determine accuracy, the results from these questionnaires will be compared to two criterion measures, the energy expenditure results recorded by the SP3 as well as one of the most widely used and available accelerometers, the GT1M. Both monitors will be worn for the same seven day period over which the subjects will be asked to complete the questionnaires. The results from this study should provide a more accurate estimate of activity levels in older adult populations, therefore allowing more accurately prescription of physical activity programs.

CHAPTER 2. REVIEW OF LITERATURE

Introduction

The many benefits of a physically active lifestyle are well established as are the detrimental effects of physical inactivity (Roberts & Barnard, 2005). However, despite the evidence of these health and lifestyle benefits, 61% of adults aged 65 and over do not meet the recommended guidelines of at least 30 minutes of moderate-intensity physical activity on 5 or more days a week (Center for Disease Control Behavioral Risk Factor Surveillance Survey, 2008). The elderly population is increasing both in number (Hetzel & Smith, 2000) and in life expectancy (78 years of age) (Arias, 2007). Therefore, increasing the prevalence of physical activity in older adults can reduce the public health burden by decreasing their risks for chronic diseases such as cardiovascular disease (Oguma & Shinoda-Tagawa, 2004) and diabetes mellitus (Knowler et al., 2002). It will also increase their quality of life and enable them to age better.

Physical activity is defined as “any bodily movement produced by skeletal muscles that results in caloric expenditure” (Schoeller & Santen, 1982). Physical activity can be placed into separate categories, such as occupational physical activity, leisure time physical activity, household chores, and transportation. In order to determine how best to increase physical activity in older adults, it is essential to ascertain their current physical activity habits. To do so, an accurate and reliable method of physical activity assessment, which specifically targets this population, is needed. There are many ways to assess physical activity, including direct observation, accelerometry, indirect calorimetry (IC), and doubly labeled water (DLW). These methods are not always practical or cost effective for larger populations. Because of their low cost, the preferred method of physical activity assessment is often physical activity questionnaires. However, relatively few instruments have been developed specifically for older adults and it is unclear which is the most accurate in this population. Therefore, the purpose of this literature review is to describe the benefits and disadvantages of select questionnaires which have been developed for the older adult population as well as a general population questionnaire which might suffice as a physical activity assessment tool for older adults. Recent research with the instruments used in this study will also be discussed.

Questionnaires

Questionnaires are a popular method of physical activity assessment due to their ease of use and low relative cost. Additional benefits of questionnaires over other methods, such as DLW, include their ability to capture the duration, intensity, and type of physical activity as opposed to just total energy expenditure (TEE) over a period of time (Welk, 2002). However, there are also some disadvantages to the use of physical activity questionnaires. Some of these include the misinterpretation of questions by the participant, inaccurate recall of activities' duration or intensity, and the resulting inaccurate estimation of TEE (Washburn, Smith, Jette, & Janney, 1993; Welk, 2002). Questionnaires also identify high-intensity activity more accurately than light or moderate-intensity activity (Bonnefoy et al., 2001; Washburn, Jette, & Janney, 1990). For example, Bonnefoy et al. (2001) administered ten questionnaires spanning various time periods (e.g. week to 3 months) to a group of healthy older men (n=19, age: 73 ± 4 years). The TEE estimates obtained from these questionnaires were then compared to those obtained using DLW and results from a VO_2 max test. Statistically significant positive correlations were seen between the questionnaires which recalled more intense activities and the DLW measurement (r = 0.18-0.65), VO_{2max} (r = 0.26-0.62), and Total Energy Expenditure/Resting Metabolic Rate (TEE/RMR) ratio (r = 0.19-0.75). Additionally, since physical activity patterns change over time, the TEE from the questionnaires which covered the same time period as the DLW correlated better than those covering just a “typical” week in the past year or a longer time period than the DLW.

Many physical activity questionnaires have been developed for younger populations (Taylor et al., 1978; Sallis et al., 1985; Dannenberg, Keller, Wilson, & Castelli, 1989). These “age-neutral” questionnaires may be inaccurate when used with an older adult population (Washburn, Jette, & Janney, 1990). One study examined the accuracy of one general population survey, the Center for Disease Control’s (CDC) Behavioral Risk Factor Surveillance Survey (BRFSS), which is usually administered via telephone to adults over the age of 18, when administered to an older adult population (Washburn, Jette, & Janney, 1990). When time spent in physical activity was compared to a physical activity diary, the BRFSS underestimated total physical activity in older adults by a mean 2 hours and 45 minutes a day. The best agreement was found between higher intensity activities, as strenuous activity was underestimated by only 5 minutes a day, while light intensity activity was underestimated by 2 hours and 20 minutes a day. While the differences in estimated and actual time spent in light intensity activity seems dramatic, this could be due to the emphasis of the BRFSS questionnaire on higher intensity physical activities, such as competitive sports, rather than lighter intensity activities. The latter compromise a majority of the activities engaged in by older adults.

Consequently, because of results from studies such as these, there have been several questionnaires developed specifically for use in the older adult population, such as the Physical Activity Scale for the Elderly (PASE) (Washburn, Smith, Jette, & Janney, 1993) and the Yale Physical Activity Survey for the Elderly (YPAS) (Richardson, Ochoa, & Wang, 2001). There are several differences in questionnaires used in younger people and those developed for the older population. These include shorter questionnaires for older adults to increase recall accuracy, more types of activities in which older adults participate in such as walking, gardening and other lighter chores, as well as fewer open-response questions (Washburn, Smith, Jette, & Janney, 1993). The reliability and validity of these questionnaires have been evaluated individually.

Physical Activity Scale for the Elderly (PASE)

The PASE was developed by Washburn and colleagues (Washburn, Smith, Jette, & Janney, 1993) based upon reviews of over 40 previous publications as well as the results of a pilot test of the questionnaire. The first draft of the questionnaire was compiled with information gathered from the studied publications by determining which information from each category was most relevant for older adults. The PASE that was developed is a short 10-item questionnaire which assesses physical activity from the past week. The questionnaire includes questions not only on occupational, household, and leisure time activities but also living situation, sleep, and restricted activity days. The frequency of these activities are classified as never, seldom (1-2 days/week), often (3-4 days/week), and mostly (5-7 days/week). Duration is also classified as less than 1 hour, between 1 and 2 hours, between 2 and 4 hours, and more than 4 hours per week. The final PASE activity score is determined by multiplying the amount of time spent in each activity (hr/week) by an item weight.

The first draft of the PASE questionnaire was pilot tested in a small sample (age 65-74, n = 12; age 75-84, n = 15; age 85+, n = 9) of older adults living around Boston and Amherst, MA. The questionnaires were administered by trained interviewers. Based upon results from the pilot test, the questionnaire was modified from its original version.

The PASE was then administered to a larger sample of older adults (n = 396). The PASE had reasonably high test-retest reliability (r = 0.75), with the mail version having a slightly higher reliability (r = 0.84) than the telephone version (r = 0.68). As expected, PASE scores declined with age (r = -0.13) and men exhibited consistently higher scores compared to women in each age group. The scores were also significantly correlated with two health status indicators, the Sickness Impact Profile (r = -0.42) and perceived health status (r = 0.34). PASE scores were positively correlated with grip strength (r = 0.37), static balance (r = 0.33), and leg strength in dominant (r = 0.25) and non-dominant (r = 0.28) legs. PASE scores also exhibited seasonal fluctuations, with higher scores in the warmer months compared to the colder months (r = 0.83). Individuals suffering from hypertension and chronic respiratory diseases scored significantly lower on the PASE. In this sample, the PASE scores were skewed slightly to the right.

Later work supported many of the original findings, such as age being negatively correlated with PASE scores (r = -0.21) as well as systolic blood pressure (r = -0.18) and positively correlated with balance (r = 0.20) (Washburn, McAuley, Katula, Mihalko, & Boileau, 1999). Scores were also significantly positively correlated with peak oxygen uptake (r = 0.20). Schuit and coworkers slightly modified the PASE to more closely match the Dutch lifestyle and compared those scores to energy expenditure assessed by DLW method over a 2 week period in Dutch elderly (Schuit, Schouten, Westerterp, & Saris, 1997). The modified version of the PASE was completed at the end of the second week. The correlation between the PASE score and Physical Activity Ratio (PAR = Total Energy Expenditure/Resting Metabolic Rate) was positive (r = 0.68).

One study examined several questionnaires aimed specifically at older adults, including the YPAS, PASE, CHAMPS, and modified Baecke, in comparison to measures of physical function (Moore et al., 2007). Participants (n = 54, men = 11, women = 43, 70% African-American) ranged from 50 to 93 years of age (mean = 67 years) and participated in a physical activity and nutrition intervention. Data were collected over three 60-minute sessions during which all questionnaires were administered verbally by a trained interviewer. Additionally, information from the Health Status Questionnaire and the mini-mental state examination (MMSE) as well as demographic information was gathered. Participants also completed the CS-PFP10, which is a group of 17 tasks of daily living which provides functional fitness scores for strength, flexibility, balance, endurance, and coordination.

The PASE showed group differences based upon race ($p<0.01$), education level ($p<0.05$), and income ($p<0.05$). African-Americans with a high school education or higher and those with annual incomes >$20,000/year had higher scores than those who did not. PASE score seemed to be one of the best indicators of CS-PFP10 scores ($r = 0.45$, $p < 0.01$). Significant correlations for subscores included PASE total score and upper body strength ($r = 0.32$, $p<0.05$), upper body flexibility ($r = 0.39$, $p<0.01$), lower body strength ($r = 0.47$, $p<0.01$), balance/coordination ($r = 0.40$, $p<0.01$), and endurance ($r = 0.45$, $p<0.01$).

Another study compared the PASE against DLW (Bonnefoy et al, 2001). The participants in this study were 19 healthy older men (mean age = 73.4 ± 4.1 years). Participants completed a VO_{2max} test at the beginning of the 2-week test period. The PASE was positively correlated with PAR ($r = 0.36$). It was also positively correlated with DLW TEE ($r = 0.28$) as well as the VO_{2max} test ($r = 0.33$).

Overall, the PASE is a brief, easily administered questionnaire which is reasonably valid (Bonnefoy et al, 2001; Schuit, Schouten, Westerterp, & Saris, 1997; Washburn, McAuley, Katula, Mihalko, & Boileau, 1999). However, the PASE measures physical activity of the previous week rather than physical activity in a typical week. This is important to remember as physical activity can be influenced by external influences, such as weather, which could cause deviation from the normal pattern (Schuit, Schouten, Westerterp, & Saris, 1997).

Yale Physical Activity Survey (YPAS)

The YPAS is an interviewer-administered, 36 question physical activity questionnaire developed in the late 1980s (DiPietro, Caspersen, Ostfeld, & Nadel, 1992). To develop this questionnaire, a total of 222 participants were drawn from urban and suburban senior centers and senior residential communities. Participants were from various socio-economic backgrounds and agreed to participate in an open-ended interview. These included questions concerning the types of activities which are most often engaged in by older adults and the reasons for participating or not participating in these activities. This information was used to formulate the YPAS.

The questionnaire assesses physical activity patterns in a “typical week” from the past month. It has two main sections from which eight scores are tabulated. The first section of the questionnaire contains a checklist with questions concerning work, exercise, and recreational activities. This assesses time spent in related activities with activity participation expressed in hours per week. The second section of the questionnaire contains categorical responses which assess participation in several different types of physical activity, including vigorous activity, low intensity activity, walking, and general movement. Three main indices can be calculated, including total time (hours/week), energy expenditure (kcal/week), and activity dimensions. In calculating total time, the time for each activity in the checklist is summed over all activities. To calculate energy expenditure, the time for each checklist activity is multiplied by an intensity code, and then summed for a total of kcal/week. Finally, the activity dimension index is determined by combining the scores of five different activities, including vigorous activity, leisurely walking, moving, standing, and sitting. The score for each of these is calculated by summing the duration of each activity and taking the duration multiplied by a frequency score. That number is then multiplied by a weighting factor, which is dependent upon the intensity of the activity performed.

After the development of the YPAS, researchers performed sub-studies which examined both the validity of the YPAS and the 2-week reliability of the YPAS. Participants for both sub-studies were subjects from the original participant pool who agreed to be recontacted. Of the original 134 who agreed to be recontacted after the initial study, 76 [12 men (71.0 ± 6.8 years) and 56 women (71.1 ± 6.3 years)] completed the repeatability study and 25 [14 men (70.7 ± 5.5 years) and 11 women (68.0 ± 5.6 years] qualified for the validity study.

For the repeatability study, participants were administered the YPAS twice by the same interviewer with 14 days between the first and second administration. Repeatability was assessed for each question of the checklist portion as well as each of the indices. The validity study included measurement of several physiologic variables, including skinfolds, height, and weight, and VO_{2max} measured via indirect calorimetry during a treadmill test. Additionally, a Caltrac accelerometer was worn for 2.5 days at hip height with Caltrac readings recorded every 3 hours in an activity diary. After the collection of all physiologic variables, participants were administered the YPAS by a trained administrator.

The YPAS was adequately repeatable and demonstrated adequate validity. For the repeatability sub-study, there were no significant differences in mean activity hours between the first (34.4 ± 18.4 hours/week, 7,613 ± 4,502 kcal/week) and second (30.6 ± 16.7 hours/week, 6,739 ± 4,267 kcal/week) administrations between any of the checklist activities. Test-retest correlations for the three main indices were 0.57 for total time, 0.58 for energy expenditure, and 0.65 for the activity dimensions. The correlations for the sub-scores of the activity dimensions index ranged from 0.42 (sitting) to 0.61 (vigorous), with higher correlations found for the more vigorous activities. As for the validation sub-study, the YPAS activity dimensions summary index was significantly associated with estimated VO2max ($r = 0.58$, $p = 0.004$) as well as percent body fat ($r = -.043$, $p = 0.03$). It was also found that the moving index was associated with resting diastolic blood pressure ($r = -0.33$, $p = 0.10$) and body mass index ($r = -0.37$, $p = 0.06$) and the sitting index correlated positively with resting diastolic pressure ($r = 0.53$, $p = 0.01$).

Other studies have also examined the accuracy and repeatability of the YPAS. One of these studies examined the 2-week repeatability and accuracy in volunteers (n = 56, age: 56-86 years) at two different sites (Schuler, Richardson, Ochoa, & Wang, 2001). Participants at both sites completed a medical history with exclusion criteria of cardiovascular disease history or elevated resting blood pressure above 160/90. Upon the first visit, the participants were administered the YPAS by a trained volunteer (Time 1). During the second, visit body fat percentage was assessed by the skinfold method and each participant completed a graded sub-maximal bicycle stress test. Two weeks after the initial administration of the YPAS, there was a second administration to assess repeatability (Time 2). At the completion of this they were asked to keep a physical activity diary for seven days. The participants were asked to write a general description of the activity, an estimation of the intensity, and the number of minutes spent in the activity. Upon completion, the diary was handed in to an investigator who reviewed the completed diary with the participant to maximize accuracy.

Repeatability for each measure of the YPAS (total time, total energy expenditure, and total activity summary index) was examined separately and showed moderate to good short-term repeatability (r = 0.70-0.82). However, the subcategories for the total activity summary index ranged from low to moderate short-term repeatability for vigorous, walking, and moving (r = 0.51-0.70) to no repeatability for either sitting or standing (r = 0.16-0.27). It is noted, however, that lower repeatability measures may not necessarily demonstrate lower repeatability but only differences in amounts of physical activity performed from week to week.

As for validity measures, no differences between males and females were seen. Statistically significant correlations were found between physical activity as noted in the physical activity diary and the total time index (r = 0.27, Time 2), total energy expenditure (r = 0.27,0 .30, Times 1 & 2), total activity summary index (r = 0.36, Time 2), and vigorous activity subcategory (r = 0.39, Time 2). Significant correlations were also found between predicted maximal oxygen consumption and total time (r = 0.49, Time 2), total activity summary index (r = 0.64, Time 2), vigorous activity (r = 0.54, Time 2), and moving (r = 0.50, Time 2). It is also important to note that greater the most significant associations were between the second administration of the YPAS (Time 2) and the validation criteria. This could be due to being more familiar with the YPAS after the first administration as well as the time of the second administration overlapping more with the keeping of the physical activity diary.

The accuracy of the YPAS has also been examined in older adults in comparison to accelerometers and DLW (Starling, Matthews, Ades, & Poehlman, 1999). Participants included healthy older men (n = 32) and women (n = 35) between the ages of 45 and 84 years. To collect data, it was required that on the first day participants stay overnight at a testing facility where they were administered their dose of DLW as well as the YPAS, in addition to one other questionnaire, by a trained interviewer. The following morning, resting metabolic rate (RMR) was determined from 45 minutes of indirect calorimetry, body composition via duel-energy X-ray absorptometry, and aerobic capacity were measured. Ten days after the initial assessment, subjects returned to provide final urine samples. During the 10 days out of the testing facility, the Caltrac accelerometer was worn during all waking hours.

For the data analyses, average daily physical activity energy expenditure over the measurement period was used from the Caltrac, DLW, and YPAS. For the women, physical activity expenditure measured by the Caltrac (379 ± 162 kcal/day) was significantly lower than the physical activity energy expenditure measured by the YPAS (863 ± 447 kcal/day) and DLW (873 ± 244 kcal/.day). For men the results were similar, with the Caltrac also showing significantly lower physical activity expenditure values (554 ± 242 kcal/day) when compared to the YPAS (1,107 ± 612 kcal/day) and DLW (1,211 ± 429 kcal/day). No difference was found between the YPAS and DLW for either men or women. However, Bland-Altman data indicate that the limits of agreement were wide for both men (-1,310 to +1,518 kcal/day) and women (-963 to +981 kcal/day). Results from this study showed that the YPAS compares favorably with DLW in group evaluation; however, uses individually may be limited in older men and women.

One study examined the accuracy of the YPAS by using the intake-balance method in assessing the energy expenditure of physical activity (PAEE) (Kruskall, Campbell, & Evans, 2004). A total of 28 participants (11 men and 17 women) with an age range from 55 to 78 years participated in the 14-week study. Weeks 2, 3, 8, and 14 were conducted on an inpatient basis while weeks 1, 4-7, and 9-13 were conducted on an outpatient basis. All subjects were sedentary during weeks 1 and 2, during which they were each assigned to one of three groups: sedentary, lower body resistance exercise or whole body resistance exercise during weeks 3-14. The YPAS was administered at weeks 2, 8, and 14 by the same interviewer each time. During the weeks of the study, all subjects consumed a controlled diet. The diet provided 0.8 grams of protein per kilogram of body weight, and non-protein energy content of 60% carbs and 40% fat to maintain body weight within ±0.5 kg of baseline body weight. All meals were consumed in the presence of a registered dietitian or diet technician.

The results showed that the women were able to maintain their baseline bodyweight while the men's declined slightly. The PAEE for men was figured at 595 ± 103 kcal/day, women at 412 ± 68 kcal/day, and for both sexes combined 486 ± 59 kcal/day. At week 14, the YPAS estimated the PAEE for women at 565 ± 93 kcal/day, men at 625 ± 93 kcal/day, and the combined estimation was 588 ± 53 kcal/day. This provides us with a statistically significant correlation ($r = 0.30$), although considered a low to moderate correlation. The results from this research project seem to support the results found by Starling and colleagues (Starling, Matthews, Ades, & Poehlman, 1999). Although the purpose of this article was not to assess repeatability, it appears to support the findings of Schuler and colleagues (Schuler, Richardson, Ochoa, & Wang, 2001) in demonstrating moderate to good repeatability.

Repeatability of the YPAS has been examined in Mexican-American older adults (Pennathur, Magham, Contreras, & Dowling, 2004). A total of 49 participants were recruited for this study (women = 42, men = 7) from senior centers in El Paso, Texas (mean age = 74 years).

Only 15 completed the second administration of the YPAS (mean age = 78 years). All participants were of Mexican origin and spent years of their childhood and young adulthood living in Mexico and lived in their own homes. There were both English and Spanish versions of the survey, as well as several bilingual interviewers who were able to administer the survey. A period of 2 weeks was allowed between the first and second administrations of the YPAS.

Results showed that activities such as shopping ($p = 0.424$), light housework ($p = 0.787$), and food preparation ($p = 0.458$) activities showed significant positive intraclass correlations. Spearman correlation coefficients showed significant positive correlations for lawn mowing ($r = 0.477$) and raking activities ($r = 0.711$). As for overall estimated energy expenditure it was significantly positively correlated ($p = 0.437$, $r = 0.626$). The total time ($r = 4.75$) and vigorous activity index ($r = 0.589$) were significantly correlated with Spearman correlation.

Previously discussed research discussed (page 5) also examined the YPAS (Moore et al., 2007). Results related to the YPAS indicate income differences with the YPAS energy expenditure summary index ($p<0.01$). The YPAS energy expenditure summary score also seemed to be a strong predictor of CS-PFP10 score ($r = 0.40$, $p<0.01$). Other significant correlations include the energy expenditure summary index and upper body strength ($r = 0.39$, $p<0.01$), lower body strength ($r = 0.42$, $p<0.01$), balance/coordination ($r = 0.34$, $p<0.05$), and endurance ($r = 0.38$, $p<0.01$).

Another study examined the YPAS as well as the 7-Day Recall in 59 individuals between the ages of 60 and 80 years and from various ethnic backgrounds (Young, Jee, & Appel, 2000). All were also participants in a 12-week study on the effects of aerobic exercise and Tai Chi on blood pressure. All participants were previously sedentary and free from any physical ailments which would prevent them from participation in an exercise program. Resting pulse rate, BMI, weight, and VO_{2max} were assessed. Each participant participated in either the aerobic or Tai Chi group program for at least one hour a day twice weekly in addition to a home-based exercise program. Between 10 and 12 weeks post-intervention start date the questionnaires were administered in random order.

Results showed that the YPAS VO_{2max} was significantly correlated with the summary index ($r = 0.33$, $P = 0.01$), moving index ($r = 0.36$, $P = 0.007$), and the standing index ($r = 0.28$, $P = 0.04$). Significant correlations were also found between BMI and the summary index ($r = -0.31$, $P = 0.02$), the moving index ($r = -0.30$, $P = 0.02$), and the standing index ($r = -0.25$, $P = 0.05$). Pulse rate was not found to be significantly correlated with any YPAS measures. As for weekly energy expenditure correlation was highest with VO_{2max} ($r = 0.20$, $P = 0.14$). Researchers concluded that the YPAS shows some promise in its ability to detect changes in physical activity.

Previously described research (page 6, Bonnefoy et al, 2001) also examined the YPAS as one of the administered questionnaires. The correlation between the YPAS TEE and DLW TEE was classified as low ($r = 0.18$). The YPAS vigorous activity index was positively correlated ($r = 0.23$) with DLW TEE. Other results show that the YPAS overestimated EEPA by around 90 kcal/day, or around 11%. Limits of agreement were -464 to 645 kcal/day for the energy expenditure score.

Overall, the YPAS seems to demonstrate good short-term repeatability (Schuler, Richardson, Ochoa, & Wang, 2001). While the repeatability seems to be good, results suggest that there is increased accuracy of the YPAS when used in assessing groups rather than assessing individuals (Kruskall, Campbell, & Evans, 2004; Starling, Matthews, Ades, & Poehlman, 1999). As with many self-report assessments, the YPAS demonstrates better accuracy for more intense physical activity (Schuler, Richardson, Ochoa, & Wang, 2001).

7-Day Physical Activity Recall

The 7-Day Physical Activity Recall (PAR) is a questionnaire developed in the general population (Sallis et al., 1985). It is a 14-item instrument in which the individual is asked to recall time spent in activity in the previous week, with a minimum duration of 10 minutes, sleeping (1 MET) and performing moderate (4 METS), hard (6 METS), or very hard (10 METS) physical activity. Time spent in light activity (1.5 METS) is calculated as [24 hours – (Time spent sleeping + time spent in moderate + hard + very hard activity)]. Total daily energy expenditure (TEE) is estimated as the sum of the amount of time (minutes/day) spent in each category of physical activity multiplied by the MET value associated with each physical activity category as well as body weight in kilograms.

The accuracy of the PAR has often been compared to the DLW method. As it was designed for use in the general population, the participants in these studies are often younger, although they still cover a wide variety of populations. One of these studies examined a total of 46 individuals (women = 29, men = 17) aged 17-35 years who were classified as overweight or moderately obese (BMI = 25.0-34.9) as well as sedentary (<500 kcal/week of leisure time physical activity). The study spanned a time period of 16 months and was aimed to assess the impact of aerobic exercise on body weight and composition. Therefore, body composition, maximal oxygen uptake, and RMR were assessed prior to the start of the study. They were then given a dose of DLW which was monitored daily for the next 14 days. The PAR was administered on day 7 or 8 of the DLW protocol to ensure that the two were overlapping.

Results showed that there were no significant differences found between the PAR (11,825 ± 1,779 kJ/day) and the DLW (11,922 ± 2,516 kJ/day) assessments for this study ($r = 0.58$, $P<0.01$). On average, the PAR underestimated daily energy expenditure (DEE) by 96 ± 2,080 kJ/day. Bland-Altman analysis showed that underestimation of DEE increased as DEE increased. Analysis of energy PAEE also showed no significant difference between the PAR (3,286 ± 502 kJ/day) and DLW (3,508 ± 1,863 kJ/day) but were not significantly correlated ($r = 0.12$). As with DEE, the underestimation of PAEE increased as PAEE increased. Results from this study suggest the PAR is able to provide an accurate estimate of DEE and PAEE but may be limited in its ability to assess individual levels.

The PAR was assessed under free living conditions in 13 healthy women (age = 21-37 years). The study protocol was seven days in length and started with an overnight stay at the research facility. Subjects consumed a dose of DLW and had RMR measured in the morning. Additionally, they wore the Tritrac-R3D and the Computer Science Applications (CSI) accelerometers during all waking hours over the next seven days. At the end of the seven days, subjects returned to the research facility to return the monitors and respond to the PAR. The PAR was interviewer-administered and the same interviewer conducted all interviews. Subjects also had their body composition assessed as well as peak oxygen consumption within four weeks of the end of the experimental period.

Results show there was no significant difference in PAEE estimates between the PAR (642 ± 35 kcal/day) and DLW (798 ± 83 kcal/day) with a mean difference of -156 kcal/day. In cases where PAEE according to DLW was less than 500 kcal/day, the PAR overestimated PAEE on average 137 kcal/day and when it was greater than 500 kcal/day, the PAR underestimated PAEE on average 287 kcal/day. Results suggest that the PAR may not accurately assess PAEE for all individuals.

Another study examined the ability of the PAR to assess EE in 14 obese women (Racette, Schoeller, & Kushner, 1994). All participants were healthy, premenopausal, nonsmokers between 21 and 47 years of age with a body fat percentage >35%. The study spanned 17 weeks and was part of a larger study to examine weight-loss treatments. Subjects were advised on a specific diet, either a low-fat or low-carb diet, to promote 1 kg/week of weight loss. Additionally, oxygen consumption and BMR measurements were performed. Subjects were then given a dose of DLW and spent one night at the research facility. Samples were also collected at the same time of day one and two weeks later. During these two weeks, subjects were asked to wear a heart rate monitor for all waking hours for at least three days during the week, as well as recording notable activities throughout the day in a folder. At the end of each two week DLW measurement period, each subject was administered the PAR.

Results showed that mean TEE values from the PAR were not statistically different than the measurements obtained from the DLW during either weight maintenance (2,665 kcal/day vs. 2,616 kcal/day, $P = 0.58$) or during the weight-loss diet (2,440 kcal/day vs. 2,452 kcal/day, $P = 0.89$). The PAR ($P = 0.92$) agreed with DLW better than the heart rate monitor ($P = 0.14$). Overall, the PAR provided accurate assessments of TEE during both weight maintenance and the diet periods and was able to detect differences in PAEE during diet periods.

Another study which examined the PAR in comparison to DLW and a physical activity diary had 27 male subjects, which were all participants in another larger feeding study (Conway, Seale, Jacobs, Irwin, & Ainsworth, 2002). Participants ranged in age from 27-65 years (mean = 41 ± 2 years). Subjects participated in a two week protocol during which DLW was administered on day one with samples being taken for each day over the next two weeks. They were also asked to keep a physical activity diary during either the first or second week with whole body calorimetry being performed on the opposite week. The PAR was administered on both day seven and 14 of the protocol. The feeding study was maintained throughout the DLW research period, with a diet of 35% fat, 15% protein, and 50% carbohydrates with a kcal intake meant to maintain body weight ± 1 kg.

Results show that the PAR overestimated free-living DEE according to DLW by 31% with a significant difference between the PAR (17.4 ± 1.45 MJ/day) and DLW (13.27 ± 0.35 MJ/day). The PAR showed higher individual variation with DLW than did the physical activity diary. Researchers in this study concluded that overestimation could be lessened by awarding lower MET values to “hard” and “very hard” physical activity categories as well as the PAR having limited abilities in assessing energy expenditure in individuals and small groups. A study previously described (Bonnefoy et al., 2001) also examined the PAR in 19 older men. Results showed that the PAR had a high strength of association with DLW ($r = 0.37$, $P<0.05$). Time spent in moderate ($r = 0.52$) , hard ($r = 0.27$), and very hard ($r = 0.18$) were all statistically significant ($P<0.05$) when compared to TEE from DLW. The PAR overestimated TEE from DLW by an average of 276 kcal/day (10.8%). On a group basis the PAR compared favorably with DLW.

One study examined the PAR in comparison with a triaxial accelerometer specifically for measuring time in exercise (Sloane, Snyder, Demark-Wahnefried, Lobach, & Kraus, 2009). Participants were also involved in a dietary study aimed to increase fruit and vegetable intake, reduce total and saturated fat intake, and increase physical activity among cancer survivors. Participants were either survivors of breast or prostate cancer. Those with any physical ailment which would prevent them from participating in an unsupervised exercise program for six months were disqualified from the study. Subjects ($n = 115$, mean age = 58 ± 10 years) reported to the research center and were provided an accelerometer to wear for seven days during all waking hours. At the end of the seven days, a version of the PAR modified to be administered via telephone, was given to each participant. Reassessments were performed after one and two years in the same manner. Results showed that the PAR was more strongly associated with the RT3 at baseline ($r = 0.54$, $P<0.0001$) and year two ($r = 0.53$, $P<0.0001$) than year one ($r = 0.24$, $P<0.01$). In all, these correlations suggest a moderate association between the RT3 and the PAR.

Previously described research (Young, Jee, & Appel, 2000) also examined the PAR in comparison to various physiological measures. Results showed that DEE as assessed by PAR ($r = 0.34$, $P = 0.01$) and moderate activity hours ($r = 0.32$, $P = 0.02$) were the only PAR measures to significantly correlate with VO_{2max}. Neither resting pulse rate nor BMI significantly correlated with any PAR measures.

Overall, the 7-day Physical Activity Recall seems to be more reliable in assessing groups and seems more limited in assessing individuals (Washburn, Jacobson, Sonko, Hill, & Donnelly, 2003; Leenders, Sherman, Nagaraja, & Kien, 2000). The PAR also seems to be more accurate in assessing those individuals with moderate activity levels as opposed to those who have extremely low (overestimate) or extremely high (underestimate) activity levels (Washburn, Jacobson, Sonko, Hill, & Donnelly, 2003). In group assessment, the PAR has demonstrated moderate to high correlations with DLW (Washburn, Jacobson, Sonko, Hill, & Donnelly, 2003; Racette, Schoeller, & Kushner, 1994). The PAR has been examined in many different populations but further examination in older adult populations is necessary.

Accelerometry-Based Activity Monitors

Accelerometers are an increasingly popular method of physical activity assessment. They are motion sensors that detect acceleration or deceleration using a piezo-electric sensor. Monitors are more sensitive in detecting motion in certain planes of motion (e.g. vertical, anteroposterior, lateral) than others. Therefore, monitors are known as uniaxial, biaxial, or triaxial based upon the number of planes in which they are most sensitive at detecting motion (Murphy, 2009). Accelerometers are most commonly worn at hip level but can be worn on other parts of the body, such as the ankle or wrist. They are often able to store large amounts of data, often days or even weeks. This information is easily downloadable for the researcher at the end of the study period. With this information, researchers are able to see objective patterns of activity throughout each day in as little as 30 second or 1-minute epochs as well as general patterns over the period of time during which the monitor is worn (Welk, 2002).

There are several benefits to using accelerometers in studies; they are easy to operate, small, non-invasive, able to store large amounts of data, and are able to provide an objective report of overall movement (Welk, 2002). However, a disadvantage of accelerometers is their inability to accurately assess various types of physical activity such as walking up an incline, carrying a load while walking, and stationary cycling (Welk, 2002). The acceleration pattern for walking up an incline and walking while carrying a load does not differ from normal walking while stationary cycling provides little or no acceleration at the hip where most accelerometers are worn (Welk, 2002). Thus, these activities are not detected as different from walking or inactivity, respectively. While accelerometers often underestimate graded walking, this may not be a serious limitation since variations in grade are not as important as variations in duration, intensity, and speed which make up the bulk of variations in daily physical activity (Montoye et al., 1983; Welk, 2002). In estimating graded walking, a triaxial accelerometer is not superior to a uniaxial accelerometer. This suggests that a triaxial accelerometer may not improve EE estimates compared to a uniaxial accelerometer (Jackic et al.,1999).

Other studies have examined the placement of the accelerometer on the body as a potential source of error. Researchers have examined the positions of accelerometers on the body, including on the ankle, wrist, waist, as well as combinations of these to examine optimal placement for these monitors. For example, one study examined the placement and orientation of accelerometers on assessment of EE during walking (Bouten, Sauren, Verduin, & Janssen, 1997). The researchers performed a complete comprehensive mechanical analysis of placement and orientation in 2 participants (age: 23-24 years). Measurements were made at the lower back using a triaxial accelerometer (ICSensors 3031-010) consisting of three separate uniaxial accelerometers with measurement directions in the sagittal, transversal, and vertical axes. The participants walked on a treadmill at 3, 4, 5, 6, and 7 km per hour for 3 minutes each. During the sessions, O_2 consumption as well as CO_2 production were monitored to predict EE. Accelerometers were placed at the lower back level with the monitors lined up on several different planes of motion. Researchers also estimated the integral of the modulus of body acceleration (IMA), as a linear relationship has previously been established, in helping predict EE. High correlations between IMA values and EE (0.96 for summed IMA from measurement directions) were found. The combined results from this study allowed researchers to conclude that monitor placement does not affect prediction of EE. While this study supported monitor placement on the lower back, the preferred site is most often the hip. This position has proven more comfortable for the participant, less obtrusive, and is well suited to measuring normal daily activity (Welk, 2002). During prolonged unsupervised periods, monitor positioning and compliance issues can be a problem as there is less guidance and supervision to ensure the monitor is being worn appropriately (Welk, 2005).

There are other limitations to the use of accelerometers. While accelerometers have proven beneficial in laboratory and field settings, the initial cost of these monitors ($200-$500) may prohibit their use in studies with large numbers of participants. Accelerometers are also unable to assess any water-based physical activity. There are also some issues in distinguishing between short periods of sedentary behavior and non-compliance with wearing the monitor, since each would result in either low or zero counts. Some monitors are sensitive enough to accumulate a small number of counts even when the participant is sedentary (Welk, 2002).

Each accelerometer has its own unit of measure, which makes it challenging to compare EE across different types of monitors. Also, EE is not assessed directly but can only be estimated when the number of counts measured by the monitor is used in an equation provided by the company or which has been developed during research (Welk, 2002). As a consequence, accelerometry-based physical activity assessment is probably more valid in assessing general physical activity behavior and less accurate in predicting EE (Welk, 2002).

GT1M Actigraph Accelerometer

The GT1M (Actigraph LLC), previously a uniaxial accelerometer known as the CSA and MTI, is now a dual-axis accelerometer. It is one of the most widely used and available accelerometers on the market. It measures and records both activity counts and steps taken. These can later be used to determine level of physical activity as well as estimate energy expenditure (EE). The GT1M is capable of recording activity in epochs ranging from one second to several minutes. When using one minute epochs, it can record up to 378 days of data (Murphy, 2009). Benefits of this version over previous versions of the accelerometer include no need for calibration, a rechargeable battery, and revisions in both hardware and software (www.actigraph.com).

Older versions of the GT1M, the CSA and the MTI, are among the most validated physical activity monitors. One meta-analysis assessed studies validating the CSA or MTI, along with other activity monitors, against DLW (Plasqui & Westerterp, 2007). The studies included in the meta-analysis assessed women, children, adolescents, Swedish, Scottish, black, and Hispanic populations with a total of 464 participants. Other uniaxial accelerometers evaluated included the Lifecorder (Suzuken Co.) and the Caltrac (Muscle Dynamics Fitness Network), while biaxial accelerometers included the Actiwatch (Minimitter Co., Inc), and triaxial accelerometers included the Tritrac-R3D (Professional Products) and the Tracmor (Philips Research). Of the uniaxial accelerometers analyzed in this study, the CSA/MTI was the only one which was repeatedly shown to correlate significantly with energy expenditure derived from DLW ($R = 0.30$-0.96). The Tracmor had the best results ($R = 0.63$-0.91), but is not available for commercial use. The results of this study are in agreement previous research which examined the CSA/MTI in comparison to the Biotrainer Pro, Tritrac-R3D, and the Actical in college-aged participants (Welk, Schaben, & Morrow, 2004). Participants in this study participated in treadmill walking at 3 mph for 5 minute bouts for each monitor. Results show that CSA/MTI was found to have the least variability across monitor units and trials and the highest overall reliability (ICC >.80).

Another study examined the ability of the GT1M, as well as several other physical activity monitors, to assess time spent in moderate and vigorous physical activity (MVPA) in comparison to indirect calorimetry (Berntsen et al., 2010). There were a total of 20 participants whom wore the monitors for a total of 120 minutes while participating in physical activities of various intensities. Results show that the GT1M overestimated time spent in MVPA by 2.5% ($p = 0.007$) but underestimated MVPA and total energy expenditure.

While several studies have been performed with the older versions of the GT1M, only a handful have examined the new version, which has been marketed since 2005 (Rothney, Apker, Song, & Chen, 2008). One study compared the performance of two previous generations of Actigraph accelerometers, including the 7164 ($n = 13$) and the 71256 ($n = 12$), to the newer version GT1M ($n = 12$) (Rothney, Apker, Song, & Chen, 2008). Motions in this study were simulated with an orbital shaker which produces accelerations similar to human walking or jogging. Five radius values, which help simulate different gaits and walking/jogging speeds, of the orbital shaker were evaluated at a consistent frequency (150 rpm). This would be similar to evaluating different subjects walking at the same pace with different stride lengths. Epoch lengths were set at one minute and measured for 6 minutes, with at least 3 minutes of rest between different radii values. The second test determined the range of the monitors. It was performed with a fixed radius value and tested at 21 different frequencies ranging from 25 to 250 rpm. Although the limit value listed for these accelerometers is around 150 rpm (2.5 Hz), 250 rpm (~4 Hz) is still within the ability of human movement. The lower limit was dictated by the ability of the machines.

The results from this study showed that all generations of Actigraph monitors had counts that were correlated with the radius at a frequency of 150 rpm (7164 $r^2 = 0.978$, 71256 $r^2 = 0.986$, GT1M $r^2 = 0.999$). The 7164 and the 71256 had similar slopes ($P = 0.229$) but the two were significantly different from the GT1M ($P < 0.001$). As for the different rpm speeds, all three monitors showed a similar curve shape. However, all the monitors were significantly different from each other at speeds over 160 rpm ($P < 0.017$). The GT1M and the 7164 were significantly different at all frequencies except 120 rpm, and the GT1M was different from the 71256 at all frequencies except for 120, 140, and 150 rpm. The 7164 and the 71256 were significantly different at 30, 35, 50, 60, 90, and 160 rpm. It was also found that the 7164 and the 71256 had similar responses to lower rpms, the GT1M required a stronger acceleration to record a non-zero reading. Finally the intermonitor variations for the GT1M (<1%) were found to be consistently lower than the 7164 or the 71256 (3-5%) monitors for frequencies greater than 40 rpm. The GT1M appears to be more consistent than previous versions of the Actigraph, but has reduced sensitivity in detecting lower frequencies as well as reduced amplitude at higher accelerations.

The validity of the GT1M in estimating EE has been assessed. One of these studies included 85 participants (age 18-70 years, 37 men, 48 women) who completed an overnight stay in a room calorimeter while three different accelerometers (GT1M, Actical, & RT3) collected activity information in one minute epochs (Rothney, Schaefer, Neumann, Choi, & Chen, 2008). Participants were also asked to complete two structured activity sessions, with the morning consisting of walking or jogging and the afternoon session consisting of sedentary activities.

Each activity was performed for 10 minutes with at least 10 minutes between activity sessions. For the remainder of the time in the room calorimeter, participants were asked to complete their normal daily activities (~15 hours). At the end of the study, the activity counts were then used with selected regression equations designed for each monitor. The GT1M was examined using three different equations known as AG 1 (Freedson, Melanson, & Sirard, 1998), AG 2 (Hendelman, Miller, Baggett, Debold, & Freedson, 2000), & AG 3 (Yngve, Nilsson, Sjostrom, & Ekelund, 2003) for the purpose of this study. All three equations were from the literature with AG 2 having been developed using mainly free-living activities.

The results from this study showed that AG 2 did not predict any time spent in sedentary (1-1.5 METs) behaviors (0.0% of total time predicted vs. 80% actual) as well as overestimated the amount of time spent in light (1.5-3 METs) activities (87.8% of time predicted vs. 16.6% actual). The AG 2 equation had the greatest percentage of error. AG 1 and AG 3 best predicted the amount of time spent in moderate activities. The AG 1 equation was better at estimating time spent in moderate (1.5% of time predicted vs. 2.1% actual) and vigorous physical activity (2.3% of time predicted vs. 1.5% actual) but under predicted mean physical activity level (1.29 predicted vs. 1.4 actual). The AG 3 equation showed no differences in physical activity level from the criterion measure (1.35 predicted vs. 1.4 actual). It was suggested by the authors that even though some of these equations seem more favorable than others, they may not be practical for assessment in free living settings. The equations assume that different types of physical activity are equally likely present in free living conditions.

Abel et al. (2008) also aimed to validate the GT1M accelerometer in walking and running in adults (Abel et al., 2008). Participants included 10 females and 10 males who completed a 10 minute walking trial during which steps were counted by two observers and indirect calorimetry was used. Six different speeds were assessed on the treadmill, ranging from 54-188 m/min. At the slowest speed, the GT1M counted 64% +/- 15% of the steps. On the remaining treadmill speeds, the GT1M undercounted all steps but by less than 3%. The Freedson equation (Freedson, Melanson, & Sirard, 1998) was used to estimate energy expenditure from the GT1M activity counts. It was found that, when compared to indirect calorimetry, the equation produced the most accurate estimates of total energy expenditure at moderate to fast walking speeds as well as moderate jogging pace.

Overall, the GT1M is one of the most valid and reliable dual-axis accelerometers commercially available. Older versions of the GT1M have produced energy expenditure values which have correlated significantly with the values produced by DLW. Finally, the GT1M has shown lower intermonitor variations (<1%) than previous versions (3-5%). Due to these reasons, it seems to be the accelerometer which is best suited for current use.

SenseWear Pro Armband 3

Recently there has been a new tool developed, the SenseWear Pro3 Armband (SP3) activity monitor, which not only estimates total energy expenditure but also assesses patterns of activity throughout the day as well as intensity, frequency, and duration. The SP3 is a lightweight armband designed to be worn over the right upper arm. It is a wireless pattern recognition monitor which assesses skin temperature, heat flux, galvanic skin responses, and motion via a 3-axis accelerometer (SenseWear, 2009). The subject's height, weight, and age are also taken into consideration for calculations by the SP3 algorithm (SenseWear, 2009).

Several studies have been performed to validate the SP3 against DLW and indirect calorimetry in several populations. Mignault et al. evaluated the SP3 against DLW in six Type II diabetic patients under free-living conditions (Mignault, St-Onge, Karelis, Allison, & Rabasa-Lhoret, 2005). Participants wore the armband for the same ten day period during which DLW assessed energy expenditure. The results showed no significant differences, with the mean daily energy expenditure (DEE) from the SP3 at 2,237 ± 568 kcal/day and DLW at 2,315 ± 625 kcal/day. The correlation between the SP3 and DLW reached r=0.97 ($P < 0.001$). St-Onge et al. (2007) also evaluated the SP3 against DLW in 45 free-living individuals aged 20-78 years (St-Onge, Mignault, Allison, & Rabasa-Lhoret, 2007). Resting metabolic rate (RMR) was estimated using indirect calorimetry and the thermic effect of food was estimated as 10% of DEE. Physical activity energy expenditure was estimated by subtracting the RMR and thermic effect of food from the DEE estimates. The results showed the mean DEE underestimation of the SP3 was 117 kcal/day (2,375 ± 366 kcal/day, $P<0.01$) but error was consistent with an intraclass correlation (ICC) of 0.81.

There have also been studies which examined the validity of the SP3 at rest. Malavolti et al. (2007) examined 99 subjects (mean age: 39 ± 15 years) at rest using the SP3 in comparison to indirect calorimetry (Malavolti et al., 2007). Participants had resting energy expenditure (REE) measured in the morning, following at least 12 hours of fasting and refraining from structured physical activity for at least 24 hours. To assess REE, participants were asked to lie quietly in an isolated room for 30-40 minutes while wearing the SP3 with VO_2 assessed via indirect calorimetry. The REE estimates from the SP3 (1540 ± 280 kcal/day) were not significantly different from the estimates obtained by indirect calorimetry (1700 ± 330 kcal/day) and were significantly correlated with one another ($r = 0.86$, $P < 0.0001$). Cereda et al. (2007) also examined the SP3's ability to assess REE in comparison to indirect calorimetry in cancer patients (Cereda et al., 2007). Ten patients participated in this study (mean age: 56.6 ± 13.3 years). REE was assessed on days 0, 7, 14, and at the end of treatment at a local hospital. REE assessments were done in the morning after fasting for 12 hours. During REE assessment, participants rested supine for 45 minutes. There were no significant differences between SP3 estimates of REE (1705 ± 278 kcal/day) and the indirect calorimetry estimates (1645 ± 282 kcal/day). The correlations between the two methods in this study was also high (r = 0.84).

The SP3 may be less adept at assessing exercise. When Fruin & Rankin evaluated the SP3 compared to indirect calorimetry in 13 aerobically untrained individuals, the SP3 was found to overestimate the energy expenditure (EE) of walking at 3mph by 14% and at 4mph by 38% (Fruin & Rankin, 2004). Walking on a 5% grade underestimated EE by 22%. No significant differences were found between the measurements of resting energy expenditure measured via indirect calorimetry when compared to estimates by the SP3 and were highly correlated (r = 0.76). Another study examined the EE during treadmill walking, cycle ergometry, arm ergometry, and stair stepping in 40 men and women aged 18-35 (Jakicic et al., 2004). Individuals participated in each exercise protocol in random order with the pace regulated by a metronome. The energy expenditure (EE) estimates from the SP3 for each exercise protocol were compared to open-circuit calorimetry. When the EE was estimated with the general algorithm supplied by the company, it underestimated the EE of walking (ICC = 0.77), cycle ergometry (ICC = 0.28), and stair stepping (ICC = 0.63) and overestimated arm ergometry (ICC = 0.74). However, when exercise-specific algorithms were applied, the estimates improved dramatically. The ICC increased for walking (ICC = 0.87), cycle ergometry (ICC = 0.89), stair stepping (ICC = 0.82), but decreased for arm ergometry (ICC = 0.66). The researchers concluded that in order to receive more accurate estimates from the SP3 exercise-specific algorithms must be applied. The algorithms for the SP3 have since been refined by the manufacturer. Another previously described study (pg. 27; Berntsen et al., 2010) also examined the SP3. Results for the SP3 show that it overestimated MVPA by 2.9%.

The SP3 has been examined with obese individuals (Papazoglou et al., 2006). The participants in this study included 142 obese individuals (BMI = 42.3 ± 7.0) as well as 25 lean and slightly overweight individuals (BMI = 25.3 ± 3.2). All participants had REE estimated in the morning after fasting for 12 hours with the SP3 against indirect calorimetry. Of the obese individuals, 29 also participated in exercise protocols. The exercise protocol included cycle ergometry, stepping, and walking. It was found that the SP3 underestimated REE in comparison to indirect calorimetry (1811 ± 346 kcal/day vs. 1880 ± 382 kcal/day) but highly overestimated EE in the obese individuals during the exercise sessions. The overestimation for cycle ergometry was 19%, stair stepping 31%, and walking 31%.

The SP3 has also been evaluated against four other accelerometers, the CSA, TriTrac-R3D, RT3, and the BioTrainer-Pro (King, Torres, Potter, Brooks, & Coleman, 2004). The participants in this study included 21 individuals with high fitness levels (VO_{2peak}>50 ml/kg/min) who walked and jogged at several different speeds. Indirect calorimetry was used as the criterion method. Even though the SP3 overestimated total EE at most speeds when compared to indirect calorimetry, it more accurately assessed EE when compared to the accelerometers. The slower speeds (54 m/min) were assessed less accurately with only a moderate correlation with indirect calorimetry (r = 0.65) than the higher speeds (80, 214 m/min) which reached and stayed at high correlations (r = 0.82). The SP3 was found to be most accurate at a moderate pace (161 m/min) (r = 0.85).

Several more recent studies have been performed further examining the reliability and validity of the SP3. One such study examined the accuracy of the SP3 in measuring DEE in healthy adults under free living conditions (Johannsen et al., 2010). This study had a total of 30 healthy adults aged 30-64 years who wore the SP3 for 14 consecutive days under free living conditions. The results obtained from the SP3 were compared to the results obtained from DLW, which was assessed over the same 14 day period. Average DEE estimates obtained from the SP3 were within 112 kcal/day of the DLW criterion method (ICC = 0.80). Another study further examined the two different software versions of the SP3 specifically in older adults (Mackey et al., 2011). A total of 19 older adults with a mean age of 82.0 years participated in this study. Participants wore the SP3 for a mean of 12.5 days (SD = ± 1.1 days), including sleeping time. Energy expenditure over the same time period was also evaluated with DLW. Results showed that there was reasonable agreement with no difference in means between DLW (2,040 ± 472 kcal/day), SP3 software version 6.1 (2,012 ± 497 kcal/day, $p = .593$), and SP3 software version 5.1 (2,066 ± 474 kcal/day, $p = .606$). Further examination shows that individual methods were highly correlated, with the SP3 software version 6.1 showing an ICC of $r = .893$ ($p<0.001$) and the SP3 software version 5.1 having an ICC of $r = .901$ ($p<0.001$).

One study has examined the validity of a previous version of the SenseWear armband, the SP2, against the IDEAA activity monitor (Welk, McClain, Eisenmann, & Wickel, 2012). Participants in this study were 30 college-aged participants and wore the SP2 and the IDEAA monitor while participating in normal activities of daily living. Results showed that the SP2 was able to provide EE estimates of within 0.10 METS of the EE estimates of the IDEAA monitor.

The SP3 has been evaluated under many different conditions, including free-living (Mignault, St-Onge, Karelis, Allison, & Rabasa-Lhoret, 2005; St-Onge, Mignaults, Allison, &

Rabasa-Lhoret, 2007) and laboratory conditions (Fruin & Rankin, 2004; Jakicic et al., 2004; King, Torres, Potter, Brooks, & Coleman, 2004), rest (Cereda et al., 2007; Malavolti et al., 2007) and exercise (Jackic et al., 2004). There have been several populations examined with the SP3, including Type II diabetic patients (Mignault, St-Onge, Karelis, Allison, & Rabasa-Lhoret, 2005), cancer patients (Cereda et al., 2007), obese individuals (Papazoglou et al., 2006), high fitness level (King, Torres, Potter, Brooks, & Coleman, 2004), aerobically untrained (Fruin & Rankin, 2004), and children (Arvidsson, Slinde, Larsson, & Hulthen, 2007). With the results from all of these projects combined, the SP3 appears to be a reasonably accurate method to accurately assess energy expenditure.

There are some advantages of the SP3 over DLW as the criterion measure in studies. The monitor is priced similar to other activity monitors and can store up to 11 days worth of data. It is non-invasive for the participant; data are easily analyzed with the correct program for the researcher. Finally, the SP3 is able to assess different activity patterns while DLW cannot.

Summary

Reliable and valid physical activity assessment questionnaires aimed at older adult populations are somewhat limited. However, with the growing population of older adults as well as the increase in life expectancy, there has been increased interest in developing methods specific to this population, including the development of several population-specific questionnaires. While the validity of these questionnaires has been examined somewhat, questions remain on how the results from these questionnaires compare not only to each other, but to an accurate criterion method. Thus, the aim of the present study is to assess the accuracy of these questionnaires and their ability to compare to each other by comparing their energy expenditure estimates to the estimates provided by the SP3 and the GT1M accelerometers.

CHAPTER 3. METHODOLOGY

Participants

Participants in this study included 36 men and women aged 60 and older. Individuals less than 60 years of age or those with an implanted electronic device were excluded from the study. In return for participation, participants were given their choice of either a $25 gift card or a free semester of membership in the Exercise Clinic at Iowa State University. Prior to any data collection, approval for the study was obtained from the Institutional Review Board of Iowa State University.

Study Protocol

Visit One

During the first meeting, participants received a thorough explanation of the purpose and methods of the study. Those still interested in participating voluntarily provided written informed consent. Weight and height were measured to the nearest tenth of a kilogram or centimeter, respectively. Participants were fitted with both the SenseWear Pro3 (SP3) and the Actigraph (GT1M) activity monitors and instructed on proper wear and positioning of each monitor. The GT1M is worn at the waistband of the right hip while the SP3 is an armband worn over the middle of the right triceps, thereby allowing them to be worn simultaneously. Participants were asked to wear each monitor simultaneously for one week, during which they were to participate in their normal activities, removing the monitors only for bathing or swimming. An appointment exactly one week later was scheduled with each participant to return their activity monitors as well as to complete the physical activity questionnaires.

Visit Two

Upon return a week later, the SP3 and GT1M were removed and data were downloaded. Next, participants completed the 7-Day Physical Activity Recall (PAR) (Sallis et al., 1985), Physical Activity Scale for the Elderly (PASE) (Washburn, Smith, Jette, & Janney, 1993), and the Yale Physical Activity Survey for the Elderly (YPAS) (Richardson, Ochoa, & Wang, 2001), which were given in a randomized order. Upon completion of the physical activity questionnaires, participants were given their choice of a $25 Target gift card or a free semester at the Iowa State University Exercise Clinic. Complete physical activity reports, including physical activity levels (PA) and energy expenditure (EE), were made available and later mailed to each participant.

Data Processing

Data recorded by the SP3 (Software version 6.1) and GT1M (ActiLife software version 4.0.4) were downloaded to a computer and analyzed with program-specific software for each monitor. Those participants not obtaining data from the monitors for at least three days for eight hours a day were excluded from the data analysis. The SP3 uses a formula based upon movement patterns along with height, weight, age, and gender to determine PAEE and total EE. Data obtained from the GT1M were applied to the ACT-F and ACT-C equations (Freedson, Melanson, & Sirard, 1998) to determine energy expenditure in kcal/day.

Questionnaires were scored by hand via their respective scales. The YPAS produced scores of both min/week as well as kcal/day, the PAR produced scores of min/week and kcal/day, while the PASE produced only min/week and an activity score. Results obtained from questionnaire scores were then compared to the criterion measure of the SP3 data, as well as the

Actigraph data. Pearson correlation coefficients were used to examine differences between both kcal/day energy expenditures as well as min/week spent in physical activity. Differences between genders were examined using one-way ANOVA. Methods showing high-correlations were examined further using paired-samples t-tests to determine differences. All methods were compared to the SP3 using Bland-Altman Plots (Altman & Bland, 1983). Comparisons were made individually between the SP3 and the alternative method to determine limits of agreement as well as the level of confidence for each.

CHAPTER 4. RESULTS

A total of 36 subjects completed data collection. Table 1 describes the characteristics of the observed population. The men were heavier and taller than the women ($P \leq 0.01$) though the BMI between the two groups did not differ.

Table 1 - Sample Characteristics (means ± SD)

	N	*Age (yr)*	*Height (cm)*	*Weight (kg)*	*BMI*
Female	28	69.7 ± 6.0	161.1 ± 5.2	66.4 ± 13.3	25.5 ± 4.5
Male	8	69.5 ± 5.5	176.3 ± 6.5**	87.3 ± 11.7**	28.1 ± 3.3
All	36	69.7 ± 5.9	164.5 ± 8.4	71.1 ± 15.6	26.1 ± 4.3

**P≤0.01 Males vs. Females

Descriptive statistics for estimates of PAEE and step count obtained from the GT1M can be found in Table 2. Average GT1M wear time for the study sample was 5.1 days and 12.3 hours/day. Statistically significant differences were found between males and females for estimates of PAEE (kcal/day) for the SP3 ($P \leq 0.05$) as well as the ACT-F PAEE (kcal/day), ACT-C PAEE (kcal/day) and step counts from the GT1M ($P \leq 0.01$). Additionally, due to various issues, data gathered with the GT1M has a slightly smaller N for females (N=21) resulting in a smaller total N (N=26) when compared to data gathered from the SP3 for both females (N=24) and total (N=29).

Table 2 - ACT EE and Step Count Characteristics (means ± SD)

	N	*SP3 PAEE (kcal/day)*	*ACT-F PAEE (kcal/day)*	*ACT-C PAEE (kcal/day)*	*SP3 Steps*	*ACT Steps*
Female	21	342 ± 260^{A}	175 ± 180	129 ± 99	8237 ± 2533^{A}	3999 ± 2947
Male	5	663 ± 392*	872 ± 474**	435 ± 65**	8292 ± 3067	10167 ± 6710**
All	26	397 ± 304^{B}	309 ± 375	188 ± 154	8247 ± 2572^{B}	5185 ± 4505

^{A}n=24, ^{B}n=29; *P≤0.05 Males vs. Females; **P≤0.01 Males vs. Females

There were robust correlations (P≤0.01) between the SP3 PAEE (kcal/day) data and physical activity estimates obtained from the GT1M for the ACT-F PAEE (kcal/day) and the ACT-C PAEE (kcal/day) equations as well as ACT Daily Step estimates (Table 3). Statistically significant correlations were also found (P≤0.05) between SP3 PA (min/week) data and ACT Daily Steps. Table 4 shows Spearman correlation coefficients. The SP3 PAEE (kcal/day) and the ACT-F PAEE (kcal/day), ACT-C PAEE (kcal/day), and the ACT Daily Steps were significantly correlated(P≤0.01). There was a high correlation between the SP3 PA (min/week) and the ACT Daily Steps which demonstrated statistical significance (P≤0.01).

Table 3 - ACT Pearson Correlation Coefficients

	ACT-F PAEE (kcal/day)	***ACT-C PAEE (kcal/day)***	***ACT Daily Steps***
SP3 PAEE (kcal/day)	.620**	.524**	.544**
SP3 PA (min/week)	--	--	.502*
SP3 Daily Steps	.384	.263	.369

**P≤0.01, *P≤0.05

Table 4 - ACT Spearman Correlation Coefficients

	ACT-F PAEE (kcal/day)	***ACT-C PAEE (kcal/day)***	***ACT Daily Steps***
SP3 PAEE (kcal/day)	.645**	.538**	.558**
SP3 PA (min/week)	--	--	.511**
SP3 Daily Steps	.492*	.382	.441*

**P≤0.01, *P≤0.05

Estimates of physical activity for individual questionnaires can be seen in Tables 5 and 6, with Table 5 showing PAEE estimates in kcal/day (questionnaire score for PASE) and Table 6 in min/week. SP3 estimates are provided for comparison. Only the SP3 demonstrated statistically significant differences in estimates between genders (P≤0.05) for kcal/day of all methods.

Table 5 - PAQ EE and Score Characteristics (means ± SD)

	N	*SP3 PAEE (kcal/day)*	*PAR PAEE (kcal/day)*	*YPAS PAEE (kcal/day)*	*PASE (score)*
Female	28	342 ± 260[A]	1293 ± 967	627 ± 539	174 ± 93
Male	8	663 ± 392[B]*	1445 ± 939	282 ± 201	150 ± 90
All	36	397 ± 304[C]	1327 ± 950	550 ± 503	169 ± 92

[A] n=24, [B] n=5, [C] n=29; *P≤0.05 Males vs. Females

Table 6 - PAQ Minutes per Week Characteristics (means ± SD)

	N	*SP3 PA (min/week)*	*PAR PA (min/week)*	*YPAS PA (min/week)*	*PASE (min/week)*
Female	28	538 ± 338[A]	1801 ± 1194	989 ± 799	888 ± 681
Male	8	795 ± 486[B]	1629 ± 1091	789 ± 1072	459 ± 592
All	36	582 ± 371[C]	1763 ± 1159	945 ± 854	793 ± 678

[A] n=24, [B] n=5, [C] n=29; *P≤0.05 Males vs. Females

Table 7 shows Pearson correlation coefficients between estimates of PAEE (kcal/day), total PA (min/week), or the score derived from the respective questionnaires, and readings obtained from the criterion measure of the SP3. Of the questionnaires, only the 7-Day Physical Activity Recall (PAR) demonstrated any statistically significant association with the criterion measure. These significant relationships were seen between PAR PAEE (kcal/day) and SP3 PAEE (kcal/day) (P≤0.05) as well as SP3 TEE (kcal/day) (P≤0.01), and PAR PA (min/week) and SP3 PA (min/week) (P≤0.05). Correlations for the YPAS ranged from -.078 to .161 while correlations for the PASE ranged from .097 to .314. When using Spearman correlation coefficients as seen in Table 8, similar relationships are seen as with Pearson correlation coefficients in Table 7. The only questionnaire to demonstrate statistically significant associations with the SP3 is once again the PAR. Estimates that demonstrated significance were between the PAR PAEE (kcal/day) and the SP3 PAEE (kcal/day) (P≤0.05) and the SP3 TEE (kcal/day) (P≤0.01). When using Spearman correlation coefficients YPAS correlations range from -.055 to .170 and PASE correlations range from .120 to .267.

Table 7 - PAQ Pearson Correlation Coefficients

	SP3 PAEE (kcal/day)	***SP3 PAEE (kcal/kg/day)***	***SP3 TEE (kcal/day)***	***SP3 PA (min/week)***
PAR PAEE (kcal/day)	.464*	.366	.556**	--
PAR PA (min/week)	--	--	--	.376*
YPAS PAEE (kcal/day)	-.039	.048	-.078	--
YPAS PA (min/week)	--	--	--	.161
PASE Score	.266	.311	.238	.314
PASE PA (min/week)	--	--	--	.097

**P≤0.01, *P≤0.05

Table 8 - PAQ Spearman Correlation Coefficients

	SP3 PAEE (kcal/day)	***SP3 PAEE (kcal/kg/day)***	***SP3 TEE (kcal/day)***	***SP3 PA (min/week)***
PAR PAEE (kcal/day)	.424*	.361	.524**	--
PAR PA (min/week)	--	--	--	.295
YPAS PAEE (kcal/day)	.045	.066	-.055	--
YPAS PA (min/week)	--	--	--	.170
PASE Score	.259	.260	.250	.267
PASE PA (min/week)	--	--	--	.120

**P<0.01, *P<0.05

Bland-Altman plots were used to allow for a more direct comparison of the methods used versus the SP3 criterion measure. Figure 1 demonstrates differences between the data gathered from the SP3 in regards to PAEE (kcal/day) and the estimated PAEE obtained from the ACT-C equations. The ACT-C formula underestimated PAE by about 184 kcal/day. There is a systematic bias in the ACT-C equations.

Figure 1. SP3 PAEE and ACT-C Differences.

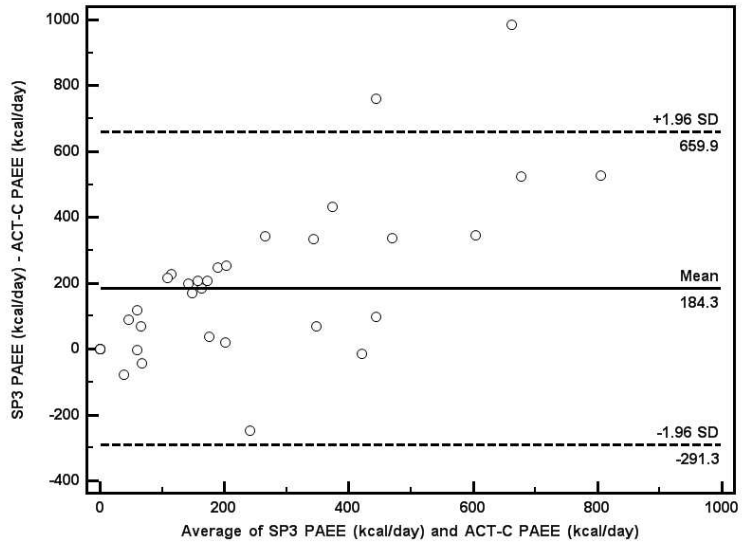


Figure 2 illustrates differences between SP3 PAEE and estimates obtained from PAR PAEE. The PAR overestimates PAEE by a mean 1007 kcal/week. The data suggest that there is a robust systematic bias.

Figure 2. SP3 PAEE and PAR PAEE Differences.

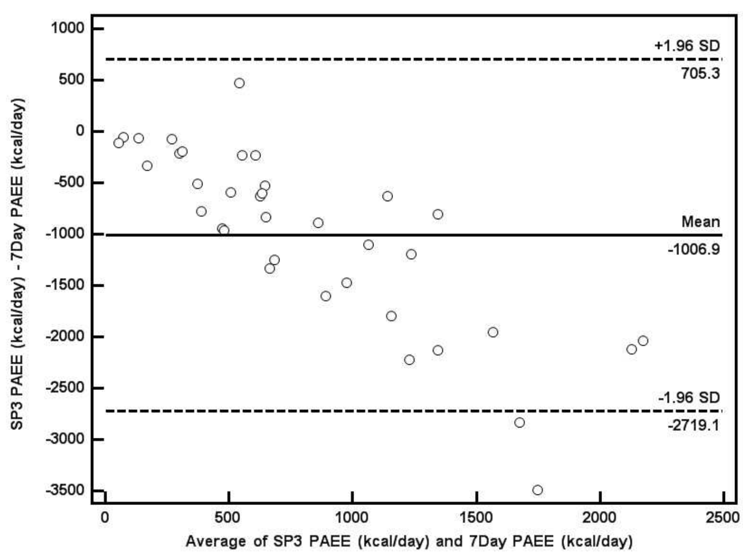


Figure 3 shows the differences between PA estimates obtained from the PAR and data obtained from the SP3. The PAR overestimates minutes spent in physical activity by about 1294 kcal/week. Again, there is a systematic bias.

Figure 3. SP3 PA and PAR PA Differences.

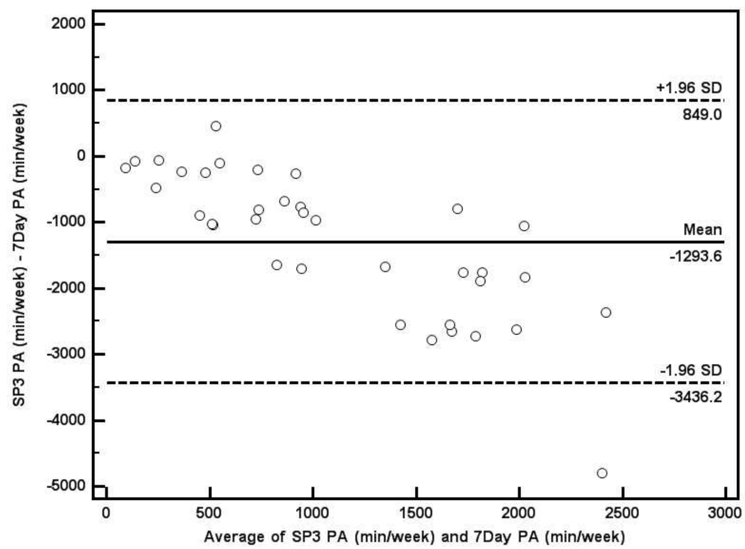
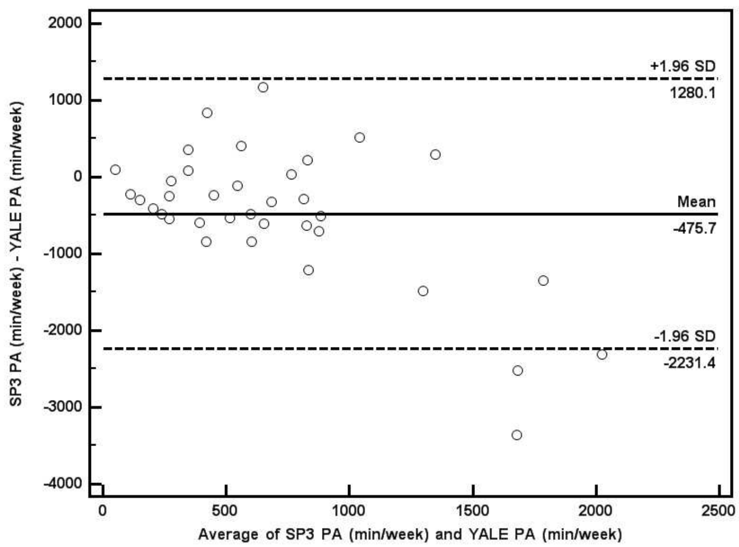


Figure 4 shows differences between the SP3 PA data and YALE PA estimates. The YALE method overestimates PA by around 476 min/week. At lower physical activity levels (<1000 min/week), the measurement bias seems to be random while at higher physical activity levels (>1000 min/week) the biased seems to be more systematic.

Figure 4. SP3 PA and YPAS PA Differences.



Differences between the SP3 measurement of PAEE and estimates in PAEE obtained from the YPAS are demonstrated in Figure 5. The YPAS survey overestimates PAEE by a mean 230 kcal/day when compared to SP3 data. These data demonstrate random bias as there seems to be no pattern in the data.

Figure 5. SP3 PAEE and YPAS PAEE Differences.

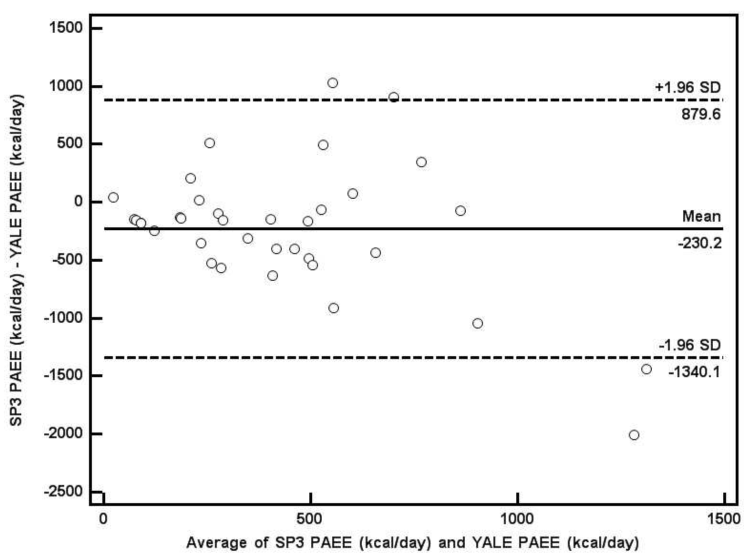
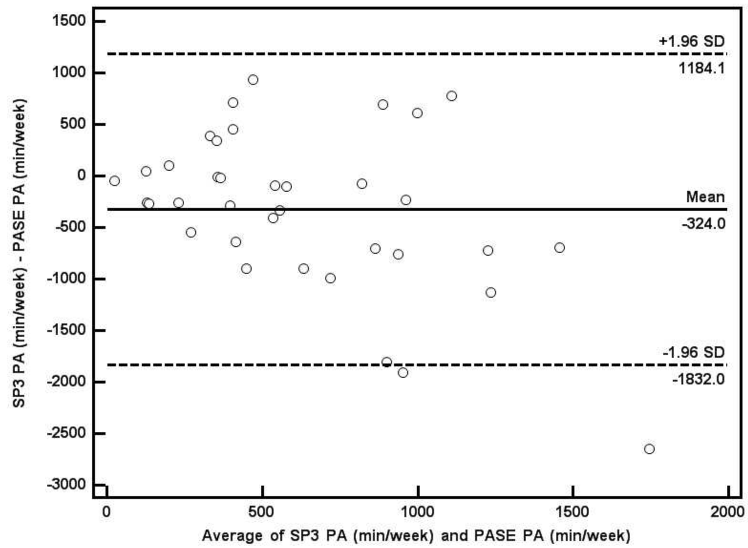


Figure 6 shows differences between the SP3 PA in min/week data and physical activity estimates obtained from the PASE questionnaire. Like other questionnaires, the PASE overestimates time spent in physical activity, this time by a mean 324 min/week.

Figure 6. SP3 PA and PASE PA Differences.



CHAPTER 5. DISCUSSION

The goal of this study was to examine both objective and subjective methods of physical activity assessment in an older adult population. This was accomplished by using three popularly used questionnaires, the 7-Day Physical Activity Recall (PAR) (Sallis et al., 1985), Physical Activity Scale for the Elderly (PASE) (Washburn, Smith, Jette, & Janney, 1993), and the Yale Physical Activity Survey for the Elderly (YPAS) (Richardson, Ochoa, & Wang, 2001), and the most popularly used accelerometer, the ActiGraph GT1M (Actigraph LLC), and comparing them to a criterion measure of the SenseWear Pro3 Armband (SP3) (SenseWear, 2009). The two monitors, the SP3 and the GT1M, were worn for one week and upon their return, participants completed the aforementioned physical activity assessments. Based upon the estimates obtained during this time period, comparisons were made between methods.

Comparison of Accelerometers

It has been suggested that an acceptable range of daily steps for healthy older adults is between 6,000-8,500 steps/day (Tudor-Locke & Myers, 2001). The daily step average seen in this study was 8,247 steps/day for the SP3 and 5,185 steps/day for the GT1M. These data fall in the recommended range for the SP3 and slightly below the range for the GT1M. Daily step counts between the SP3 and the GT1M were not significantly correlated when using Pearson correlation coefficients but were significant at the $P \leq 0.05$ level when using Spearman Correlations.

Daily steps were not the only method of comparison between the GT1M and the SP3 as kcal/day of physical activity energy expenditure (PAEE) and min/week spent in physical activity (PA) were also examined. As the GT1M does not provide direct readings of kcal/day estimates, the activity counts provided by the GT1M must be used in specialized regression equations which provide these estimates. There are several equations which can provide kcal/day estimates, two of these being the Freedson equation (ACT-F) equation (Freedson, Melanson, & Sirard, 1998) and the Combination equation (ACT-C) that were examined in this study. The ACT-F equation requires 1,952 counts per minute for that minute to be considered physically active. This equation is considered to be moderately accurate for estimating the EE of locomotor activities (Crouter, 2006). However, most activities of daily living include both locomotor activities and lifestyle activities. This is where the ACT-C equation is considered. It uses the Freedson equation for minutes with counts over 1,952 counts and the Work-Energy-Theorem for minutes with counts below 1,952 counts. Using Pearson correlation coefficients both the ACT-F equation ($r=.620$) and the ACT-C equation ($r=.524$) produced statistically significant correlations with SP3 PAEE (kcal/day) at the $P \leq 0.01$ level. Upon examination of the Bland-Altman plot for the ACT-C there is evidence of a strong systematic bias with the ACT-C underestimating physical activity expenditure by a mean 185 kcal/day. This effect is increased at higher activity levels.

Some issues with the GT1M, as with most accelerometers, are its ability to capture physical activity performed by the upper body. As most accelerometers are worn on the waist band, most are unable to capture activity that is performed when the individual is non-mobile. In the older adult population, this constitutes a majority of their physical activity, including household activities and occupational activities. This may help to explain the difference in kcal/day estimates provided by the ACT-F and ACT-C equations and the kcal/day data provided by the SP3. The SP3 is worn on the upper arm and would be able to more accurately estimate this upper body activity while the GT1M cannot. However, with the difference in body positioning it is more likely that the GT1M is more able to accurately assess step count as the

SP3 worn on the upper arm is more likely to misinterpret upper body movement as steps than the GT1M is.

Comparison of PAQs with Criterion Measure

Of the physical activity questionnaires (PAQs) examined, the PAR demonstrated the highest correlations with the criterion measure. It was the only PAQ to demonstrate statistically significant Pearson correlation coefficients with the criterion measure and did so for both PAR PAEE and PAR PA. However, further examination of the Bland-Altman plots demonstrates a strong systematic bias seen in the PAR when compared to the SP3 and shows the PAR overestimates PA more at higher levels. This is in contrast to previous studies which have found the PAR to underestimate PA in adults (Matthews & Freedson, 1995). While the PAR was not designed for use in older adults, it demonstrated the strongest correlations with the criterion measure. Similar findings between the PAR and the SP3 were found in a companion study (McIntyre, 2010).

The PASE and the YPAS were the two PAQs that were examined in this study which were specifically designed for use in older adults. Neither of them demonstrated statistically significant Pearson correlation coefficients with SP3; however, in comparison to the PAR (1,294 min/week), both the YPAS (476 min/week) and PASE (324 min/week) overestimated minutes spent in physical activity expenditure by far less. When examining PAEE in kcal/week the YPAS also overestimated by far less than the PAR with the YPAS overestimating by 230 kcal/week in comparison to the 1007 kcal/week overestimated by the PAR. A comparison of PAEE in kcal/week was not possible as the PASE did not provide this information.

There are several possibilities for the overestimation of the PARs. One of these possibilities includes “double reporting” of physical activities. For example, in the YPAS an individual could double report time spent doing dishes under the categories of light housework and dish washing, not realizing they were doing so.

Another possibility in the overestimation of the PARs is related to the relationship between absolute metabolic equivalents (METS) and relative METS. The definition of a MET is the amount of oxygen consumed while sitting at rest per kilogram of body weight and is figured at 3.5 ml 0_2/kg/min (Jette, Sidney, & Blumchen, 1990). This is known as absolute METS. Therefore individuals working at 3 METS would be working 3 times as hard as they would be at rest. However a problem with this standardization of MET activities in determining intensity of physical activities is that not every individual has a MET equivalent of 3.5 ml 0_2/kg/min, which is only able to be determined through such methods as indirect calorimetry. The actual MET levels of individuals are known as relative METS. Relative MET levels are affected by the individual’s fitness level and factors such as age, as many factors allowing a higher functional capacity deteriorate with age (Jette, Sidney, & Blumchen, 1990). Functional capacity is the maximum MET level an individual is able to obtain. Therefore individuals with lower relative MET levels and lower functional capacity MET levels would be more likely to report activities performed as more vigorous or more time was spent in vigorous activities (such as on the PAR) in comparison to the activities’ absolute MET level.

Limitations

Limitations to this study include the study sample, from which findings would be difficult to generalize to the entire older adult population. In this study there was an imbalance of men to women, with 28 of 36 total participants being female. Additionally, all study participants were

Caucasian and from southeastern Iowa. Improvements would be to take a larger sample that is more representative of the entire older adult population.

Another common issue when working with older adults includes recall bias and difficulty accurately recalling activities performed (Washburn, Smith, Jette, & Janney, 1993; Welk, 2002). While participants were able to ask questions, the best was done not to lead them. It is unknown whether this influenced their answers in any way. Additionally it is possible that the fact that all data were gathered during the autumn season could skew the data although how is unclear.

Other issues included compliance and monitor problems with the GT1M as data were able to be obtained and used from only 26 out of 36 participants. The most likely explanations include that, for various reasons, monitors stopped recording data at some point during the monitoring period or that individuals decided against and/or forgot to wear their monitors while they were ambulatory. Finally, when administering the PAQs they were inadvertently mostly given in the same order so future studies examining these PAQs should make effort to further randomize them.

Conclusion

The purpose of this study as it is stated was to further examine and validate specific PAQs designed for the older adult population, the YPAS and the PASE, in comparison to another PAQ designed for use in the general population, the PAR, and a commonly used and available accelerometer, the GT1M, was accomplished. The GT1M is one of the most commonly used accelerometers on the market and previous studies have shown it to be an accelerometer with a very low intermonitor variability (Rothney, Apker, Song, & Chen, 2008) and displayed the highest significant correlations of a group of accelerometers examined in comparison to doubly labeled water (DLW) (Plasqui & Westerterp, 2007). However in the current study the GT1M differs from our criterion measure in step count but demonstrates higher correlations when the ACT-F equation is used to estimate PAEE and PA.

When examining how the PAQs fared against the criterion measure, the PAR demonstrated the highest correlations but demonstrated a robust systematic biased. Both the PASE and the YPAS overestimated the PAEE and PA in comparison to the criterion measure, however the PASE overestimated by the least at 324 min/week. Of the two PAQs that estimated kcal/week, the YPAS overestimated by far less than did the PAR (230 kcal/week vs. 1007 kcal/week); however neither the YPAS nor the PASE demonstrated any significant correlations with the criterion measure.

When examining the accuracy of PAQs, there are many sources of error that should be controlled in order to obtain the most accurate estimates of PAEE (kcal/day) and PA (min/week) from these methods of assessment. While the best attempts can be made to control these sources of error, such as order of questionnaires, not leading participants, and obtaining a study sample representative of the population, other sources of error such as recall bias, errors in time estimates by the participants, and double reporting of activities are less able to be controlled by the researchers. The estimate errors of these PAQs are most likely not a result of any one source of error, instead, a combination of all sources of error. Learning to control each source of error individually will likely contribute to a reduction in the total error of the instrument; however, this theory should be further examined.

With the information obtained from this study, we are aware that there are many tools that can be used to assess physical activity but few have been designed specifically for use in older adults. However, as the methods examined in this study demonstrate weaknesses and lack in accuracy, they should all be used in caution in the older adult population. Results from this study agree with previous research suggesting the PASE and YPAS seem to be more accurate when used with individuals with lower PA levels (Haranda, Chiu, King, & Stewart, 2001) However, the PASE and YPAS can still be useful with individuals with higher PA levels in determining PA patterns as determining PA patterns can help in determining health status and function (DiPietro, 2001). When compared to the PAQs examined in this study the GT1M had higher correlations than did any of the PAQs. These results seem to agree with results obtained from other studies in that it correlates reasonably with a criterion method (Plasqui & Westerterp, 2007) and demonstrates higher correlations than both the PASE and YPAS with a criterion method (Colbert, Matthews, Havighurst, Kyungmann, & Schoeller, 2011) but, as with the PAQs, should be used with caution and may not always be practical due to their higher initial costs when compared to PAQs.

REFERENCES

Altman D.G., Bland J.M. (1983). Measurement in medicine: the analysis of method comparison studies. *Statistician, 32,* 307–317.

American Collge of Sports Medicine Position Stand. (1998). Exercise and physical activity for older adults. *Medicine & Science in Sports & Exercise, 30,* 992-1008.

Arias, E. (2007). National and vital statistics report, United States life tables 2004. Available at: http://www.cdc.gov/nchs/data/nvsr/nvsr56/nvsr56_09.pdf. Accessed February 17, 2009.

Arvidsson, D., Slinde, F., Larsson, S., & Hulthen L. (2007). Energy cost of physical activities in children: Validation of SenseWear Armband. *Medicine & Science in Sports & Exercise, 39,* 2076-2084.

Bernsten, S., Hageberg, R., Aandstad, A., Mowinckel, P., Anderssen, S.A., Carlsen, K.H., & Andersen, L.B. (2010). Validity of physical activity monitors in adults participating in free-living activities. *British Journal of Sports Medicine, 44,* 657-664.

Bonnefoy, M., Normand, S., Pachiaudi, C., Lacour, J.R., Laville, M., & Kostka, T. (2001). Simultaneous validation of ten physical activity questionnaires in older men: A doubly labeled water study. *Journal of the American Geriatric Society, 49,* 28-35.

Bouten, C.V.C., Sauren, A.A.H.J., Merduin, M., & Janssen, J.D. (1997). *Medical & Biological Engineering & Computing, 35,* 50-56.

Caspersen, C.J., Powell, K.E., & Christenson, G.M. (1985). Physical activity, exercise, and physical fitness: Definitions and distinctions for health-related research. *Public Health Report, 100,* 126-131.

Centers for Disease Control Behavioral Risk Factor Surveillance System; Prevalence for Data: Physical Activity, Nationwide. (2007). Available at: http://apps.nccd.cdc.gov/brfss/age.asp?yr=2007&state=US&qkey=4418&grp=0. Accessed February 17, 2009.

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

Colbert, L.H., Matthews, C.E., Havighurst, T.C., Kyungmann, K., & Schoeller, D.A. (2011). Comparative validity of physical activity measures in older adults. *Medicine & Science in Sports & Exercise, 43,* 867-876.

Conway, J.M., Seale, J.L., Jacobs, D.R., Irwin, M.L., & Ainsworth, B.E. (2002). Comparison of energy expenditure estimates from doubly labeled water, a physical activity questionnaire, and physical activity records. *The American Journal of Clinical Nutrition, 75,* 519-525.

Crouter, S.E., Churilla, J.R., & Bassett, D.R. Jr. (2006). Estimating energy expenditure using accelerometers. *European Journal of Applied Physiology, 98,* 601-12.

Dannenberg, A.L., Keller, J.B., Wilson, P.W.F., & Castelli, W.P. (1989). Leisure time physical activity in the Framingham offspring study: Description, seasonal variation and risk factor correlates. *American Journal of Epidemiology, 129,* 76-88.

DiPietro, L. (2001). Physical activity in aging: Changes in patterns and their relationship to health and function. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 56,* 13-22.

DiPietro, L., Caspersen, C.J., Ostfeld, A.M., & Nadel, E.R. (1993). A survey for assessing physical activity among older adults. *Medicine & Science in Sports & Exercise, 25,* 628-642.

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

Fruin, M.L., & Rankin, J.W. (2004). Validity of a multi-sensor armband in estimating rest and exercise energy expenditure. *Medicine & Science in Sports & Exercise, 36,* 1063-1069.

Hendelman, D., Miller, K., Baggett, C., Debold, E., & Freedson, P. (2000). Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Medicine & Science in Sports & Exercise, 32,* 442-449.

Haranda, N. D., Chiu,V., King, A.C., & Stewart, A.L. (2001). An evaluation of three self-report physical activity instruments for older adults. *Medicine & Science in Sports & Exercise, 33,* 962-970.

Hetzel, L., & Smith, A. (2001). The 65 years and over population: 2000. Available at: http://www.census.gov/prod/2001pubs/c2kbr01-10.pdf. Accessed February 17, 2009.

Hunter, G.R., McCarthy, J.P., & Barnman M.M. (2004). Effects of resistance training on older adults. *Sports Medicine, 34,* 329-348.

Jakicic, J.M., Marcus, M., Gallagher, K.I., Randall, C., Thomas, E., Goss, F.L., & Robertson, R.J. (2004). Evaluation of the SenseWear Pro Armband ™ to assess energy expenditure during exercise. *Medicine & Science in Sports & Exercise*, *36*, 897-904.

Jakicic, J.M., Winters, C., Lagally, K., Ho, J., Robertson, R.J., & Wing, R.R. (1999). The accuracy of the TriTrac-R3D accelerometer to estimate energy expenditure. *Medicine & Science in Sports & Exercise, 31,* 747-754.

Jette, M., Sidney, K., & Blumchen, G. (1990). Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity. *Clinical Cardiology*, *13,* 555-565.

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 & Science in Sports & Exercise, 42*, 2134-2140.

King, G.A., Torres, N., Potter, C., Brooks, T.J., & Coleman, K.J. (2004). Comparison of activity monitors to estimate energy cost of treadmill exercise. *Medicine & Science in Sports & Exercise*, *36*, 1244-1251.

Knowler, W.C., Barrett-Connor, E., Fowler, S.E., Hamman, R.F., Lachin, J.M., Walker, E.A., et al. (2002). Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *New England Journal of Medicine*, *346,* 393-403.

Kohrt, W.M., Bloomfield, S.A., Little, K.D., Nelson, M.E., Yingling, V.R., & American Collegeof Sports Medicine. (2004). American College of Sports Medicine Position Stand: Physical activity and bone health. *Medicine & Science in Sports & Exercise*, *36*, 1985-1996.

Kruskall, L.J., Campbell, W.W., & Evans, W.J. (2004). The Yale Physical Activity Survey for older adults: Predictions in the energy expenditure due to physical activity. *Journal of the American Dietetic Association*, *104*, 1251-1257.

Leenders, N.Y.J.M., Sherman, W.M., Nagaraja, H.N., & Kien, C.L. (2001). Evaluation of methods to assess physical activity in free-living conditions. *Medicine & Science in Sports & Exercise, 33,* 1233-1240.

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, 66,* 1108-1113.

Malavolti, M., Pietrobelli, A., Dugoni, M., Poli, M., Romagnoli, E., De Cristofaro, P., & Battistini, N.C. (2007). A new device for measuring resting energy expenditure (REE) in healthy subjects. *Nutrition, Metabolism & Cardiovascular Diseases, 17,* 338-343.

Matthews C.E., & Freedson P.S. (1995). Field trial of a three-dimensional activity monitor: comparison with self-report. *Medicine and Science in Sports and Exercise, 27*, 1071-8.

McIntyre, Patrick. (2010). *Methods of physical activity assessment for older adults*. M.S. Thesis, Iowa State University: USA.

Mignault, D., St-Onge, M., Karelis, A.D., Allison, D.B., & Rabasa-Lhoret, R. (2005). Evaluation of the portable HealthWear armband. *Diabetes Care, 28,* 225-227.

Montoye, H.J., Washburn, R.A., Servais, S., Ertl, A., Webster, J.G., & Nagle, F.J. (1983). Estimation of energy expenditure by a portable accelerometer. *Medicine & Science in Sports & Exercise, 15,* 403-407.

Moore, D.S., Ellis, R., Allen, P.E., Cherry, K.E., Monroe, P.A., O'Neil, C.E., & Wood, R.H. (2007). Construct validation of physical activity surveys in culturally diverse older adults: A comparison of four commonly used questionnaires. *Research Quarterly for Exercise and Sport, 79,* 42-50.

Murphy, S.L. (2009). Review of physical activity measurement using accelerometers in older adults: Considerations for research design and conduct. *Preventative Medicine, 28,* 108-114.

Oguma, Y., Shinoda-Tagawa, T. (2004). Physical activity decreases cardiovascular disease risk in women: Review and meta-analysis. *American Journal of Preventative Medicine, 26,*407-418.

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

Pennathur, A., Magham, R., Contreras, L.R., & Dowling, W. (2004). Test-retest reliability of Yale Physical Activity Survey among older Mexican American adults: A pilot investigation. *Experimental Aging Research, 30,* 291-303.

Racette, S.B., Schoeller, D.A., & Kushner, R.F. (1994). Comparison of heart rate and physical activity recall with doubly labeled water in obese women. *Medicine & Science in Sports & Exercise, 27,* 126-133.

Richardson, M.T., Ochoa, P., & Wang, M.Q. (2001). Accuracy and repeatability of the Yale physical activity survey in assessing physical activity of older adults. *Perceptual & Motor Skills, 93*, 163-177.

Roberts, C.K., & Barnard, R.J. (2005). Effects of exercise and diet on chronic disease. *Journal of Applied Physiology, 98*, 3-30.

Rothney, M.P., Apker, G.A., song, Y., & Chen, K.Y. (2008). Comparing the performance of three generations of ActiGraph Accelerometers. *Journal of Applied Physiology, 105,* 1091-1097.

Sallis, J.F., Haskell, W.L., Wood, P.D., Fortmann, S.P., Rogers, T., Blair, S.M., & Paffenbarger, R.S., Jr. (1985). Physical activity assessment methodology in the five city project. *American Journal of Epidemiology, 121,* 91-106.

Schoeller, D., & Santen, E.V. (1982). Measurement of energy expenditure in humans by doubly labeled water method. *Journal of Applied Physiology, 52*, 955-959.

Schuler, P.B., Richardson, M.T., Ochoa, P., & Wang, M.Q. (2001). Accuracy and repeatability of the Yale physical activity survey in assessing physical activity of older adults. *Perceptual & Motor Skills, 93,* 163-177.

Schuit, A.J., Schouten, E.G., Westerterp, K.R., Saris, W.H.M. (1997). Validity of the physical activity scale for the elderly (PASE): According to energy expenditure assessed by the doubly labeled water method. *Journal of Clinical Epidemiology, 50*; 541-546.

SenseWear: Core technology. http://www.sensewear.com/how_core.php. Accessed March 24, 2009.

Sloane, R., Snyder, D.C., Demark-Wahnefried, W., Lobach, D., & Kraus, W.E. (2009). Comparing the 7-day physical activity recall with a triaxial accelerometer for measuring time in exercise. *Medicine & Science in Sports & Exercise, 41,* 1334-1340.

Starling, R.D., Matthews, D.E., Ades, P.A., & Poehlman, E.T. (1999). Assessment of physical activity in older individuals: a doubly labeled water study. *Journal of Applied Physiology*, *86,* 2090-2096.

St-Onge, M., Mignault, D., Allison, D.B., & Rabasa-Lhoret, R. (2007). Evaluation of a portable device to measure daily energy expenditure in free-living adults. *The American Journal of Clinical Nutrition, 85,* 742-749.

Taylor, H.L., Jacobs, D.R., Schucker, B., Knudsen, J., Leon, A.S., & Debacker, G. (1978). A questionnaire for assessment of leisure time physical activities. *Journal of Chronic Diseases*, *31*, 741-755.

Tudor-Locke, C.E & Myers, A.M. (2001). Methodological considerations for researchers and practitioners using pedometers to measure physical (ambulatory) activity. *Research Quarterly for Exercise and Sport*, *72,* 1-12.

Washburn, R.A. (2000). Assessment of physical activity in older adults. *Research Quarterly for Exercise and Sport, 71,* 79-88.

Washburn, R.A., Jacobsen, D.J., Sonko, B.J., Hill, J.O., & Donnelly, J.E. (2003). The validity of the Stanford seven-day physical activity recall in young adults. *Medicine & Science in Sports & Exercise, 35*, 1374-1380.

Washburn, R.A., Jette, A.M., & Janney, C.A. (1990). Using age-neutral physical activity questionnaires in research with the elderly. *Journal of Aging & Health*, *2*, 341-356.

Washburn, R.A., McAuley, E., Katula, J., Mihalko, S.L., & Boileau, R.A. (1999). The physical activity scale for the elderly (PASE): Evidence for validity. *Journal of Clinical Epidemiology*, *52*, 643-651.

Washburn, R.A., Smith, K.W., Jette, A.M., & Janney, C.A. (1993). The physical activity scale for the elderly (PASE): Development and evaluation. *Journal of Clinical Epidemiology*, *46*, 153-162.

Welk, G.J. (2002). Physical activity assessments for health-related research. Human Kinetics Publishers,Inc.

Welk, G.J., McClain, J.J., Eisenmann, J.C., & Wickel, E.E. (2007). Field validation of the MTI Actigraph and BodyMedia armband monitor using the IDEEA monitor. *Obesity, 15,* 918-928.

Welk, G.J., Schaben, J.A., Morrow Jr., J.R. (2004). Reliability of accelerometry-based activity monitors: a generalizability study. *Medicine and Science in Sports and Exercise, 36*, 1637–1645.

Wendel-Vos, G.C., Schuit, A.J., Feskens, E.J.M., Boshuizen, H.C., Verschuren, W.M.M., Saris, W.H.M., et al. (2004). Physical activity and stroke: A meta-analysis of observational data. *International Journal of Epidemiology*, *33*, 787-798.

Yngve, A., Nilsson, A., Sjostrom, M., & Ekelund, U. (2003). Effect of monitor placement and of activity setting on the MTI accelerometer output. *Medicine & Science in Sports & Exercise*, *35,* 320-326.

Young, D.R., Jee, S.H., & Appel, L.J. (2000). A comparison of the Yale physical activity survey with other physical activity measures. *Medicine & Science in Sports & Exercise*, *33*, 955-961.