ARCHIVES
of Exercise in Health and Disease

DOI: 10.5628/aehd.v5i1-2.167

*Review Manuscript*

# Measurement of Physical Activity for Health Promotion and Education Research

**Knowlden A**

College of Human Environmental Sciences, Department of Health Sciences, University of Alabama, USA

***Objectives:*** Accurate and reliable measurement of physical activity (PA) is critical for accuracy of health education and promotion research and evaluation. In terms of research, PA measurement allows researchers to model the correlates and determinants of PA. From an evaluation perspective, measurement allows researchers to gauge the efficacy of interventions designed to increase or sustain PA. A variety of direct and indirect PA monitoring methods are available to researchers and practitioners. The purpose of this research was to review methods of measuring physical activity over time and identify best practices for application. ***Methods:*** A total of 21 articles were extracted from Medline, CINHAL, and ERIC databases for this review. Monitoring techniques analyzed included doubly labeled water, indirect calorimetry, motion sensors, global positioning system, global geographic information systems, heart rate monitors, direct observation, activity logs, and self-report questionnaires. Strengths and limitations, application, cost of the data collection, and the accuracy of the resulting data was elucidated. ***Results:*** PA is an important domain of health promotion and health education research. Accurate and reliable measurement of PA is critical to increasing the internal validity of health education interventions. ***Conclusions:*** Ultimately, the research questions that underlie a research project should determine the measurement tool researchers apply. In making a final decision, expertise, resources, and funding of the program must be considered. Even the most advanced monitoring tools have limitations. Researchers must be cognizant of the limitations of the measurement techniques they apply and make efforts to reduce biases associated with their selected measurement method.

*Arch Exerc Health Dis 5 (1-2): 338-345, 2015*

**Key Words:** Measurement; health promotion; physical activity; internal validity; pedometer; accelerometer

## ISSUES

Physical activity (PA) is recognized as a modifiable determinant in the prevention of non-communicable diseases such as cardiovascular disease, hypertension, diabetes mellitus, and obesity (26). In their seminal paper, Caspersen, Powell, and Christenson defined PA as any bodily movement produced by the skeletal muscles that result in energy expenditure (9). The amount of energy required for an activity is typically quantified through the kilocalorie, a measure of heat. Total caloric expenditure associated with PA is determined by the amount of muscle mass producing bodily movements, as well as the intensity, duration, and frequency of muscular contractions.

In 2008, the United States Department of Health and Human Services (26) released the Physical Activity Guidelines for Americans (PAGA). This focused on the preventative effects of PA on reducing the risk for chronic diseases such as heart disease and diabetes mellitus. The PAGA often serves as the reference point for the development of health promotion and physical activity interventions in education. In the document, PA is defined as bodily movement that enhances health. In this context, bodily movement is categorized as either baseline activity or health-enhancing PA. Baseline activity refers to the light-



**Corresponding author:**
Adam Knowlden: College of Human Environmental Sciences, Department of Health Science, University of Alabama, USA • Email: aknowlden@ches.ua.edu

intensity activities of daily life such as standing, walking slowly, and lifting lightweight objects. Individuals who only perform baseline activity are considered inactive. Intermittent bursts of moderate or vigorously intense activity, such as climbing a few flights of stairs, are not considered sufficient to meet the PAGA standard of health-enhancing PA (26). Conversely, health-enhancing PA produces health benefits. Examples include brisk walking, dancing, lifting weights, climbing on recess equipment, and performing yoga.

The health benefits of PA are related to the energy cost associated with a given activity (26). Physiologically, the absolute intensity of PA can be measured in terms of its metabolic equivalent (MET). Intensity of PA can also be gauged through relative intensity. Relative intensity can be expressed in terms of an individual's maximal heart rate, heart rate reserve, or aerobic capacity reserve. To simplify the scientific standards for gauging absolute and relative intensity, the PGSA has divided aerobic PA into four classifications: inactive, low, medium, and high. The classifications are based on the complementary health benefits associated with the relative levels of activity.

The inactive level includes only baseline activity (26). Minimal health benefits are associated with inactivity, and there is great evidence that inactivity contributes to numerous deleterious health outcomes. Low levels of PA include activity beyond baseline but fewer than 150 minutes of moderate-intensity physical activity per week. Minimal health benefits are associated with low level PA. Medium PA includes 150 to 300 minutes of moderate-intensity physical activity per week. The medium PA benchmark can also be met by achieving 75 to 150 minutes of vigorously-intense PA. High PA includes greater than 300 minutes of PA per week. The medium PA standards achieve the absolute intensity requirements of 500 to 1,000 MET minutes per week (26). Substantial health benefits are associated with the medium PA threshold. The medium classification of PA generally serves as the desired outcome of health promotion and education PA interventions. In integrating medium PA into interventions, it is often recommended that adults should accumulate 30 minutes of moderate-intensity PA on five or more days a week, for a total of at least 150 minutes per week.

## DESCRIPTION

PA is an important domain of health promotion and health education research. Accurate and reliable measurement of PA is critical to increasing the efficacy of health education interventions (26). Tracking PA over time can be a difficult endeavor, and evaluators conducting field-based PA research are often faced with balancing rigidity with practicality. The purpose of this study was to provide a narrative review of methods for measuring physical activity over time, elucidate each method's strengths and limitations, and identify best practices for application in health promotion and education research. Measurement techniques analyzed for this review included doubly labeled water, indirect calorimetry, motion sensors, global positioning systems, global geographic information systems, heart rate monitors, direct observation, activity logs, and self-report questionnaires.

***Lessons Learned. Direct monitoring.*** Direct monitoring measures PA through an instrument or direct observation. Direct methods for monitoring PA offer several advantages. The primary advantage is the objective measurement of PA and the elimination of recall and response biases. Although direct methods offer more robust estimates of PA, they may require in-depth training and/or the use of complex instruments which may not be practical for field research (3). Interventions employing direct monitoring also tend to require significant funding, which may prohibit their application (14).

***Doubly labeled water and indirect calorimetry:*** Doubly labeled water is the current gold standard for measuring free-living energy expenditure in humans (1). Doubly labeled water is a biochemical procedure that involves the ingestion of a quantity of water labeled with a known concentration of naturally occurring, stable isotopes of hydrogen and oxygen. Carbon dioxide and water are produced as energy is expended in the body, and the difference between the isotope elimination rates is used to calculate total energy expenditure. The standard procedure for doubly labeled water is a collection of a baseline sample before drinking the water, an initial sample after the water has equilibrated with the body water, and a final sample one to four weeks later. This duration of time can capture habitual energy expenditure patterns.

While doubly labeled water is the ideal, the cost of materials and the expertise required to analyze the isotope concentrations with mass spectrometry prohibits wide-scale application of the technique for intervention research. The technique is also rather invasive, and there are issues associated with doubly labeled water and assessing patterns of physical activity. Indirect calorimetry is increasingly being used alongside doubly labeled water to provide estimates of energy expenditure associated with PA. Indirect calorimetry, a direct monitoring technique, uses respiratory gas analysis to measure energy expenditure (14). Indirect calorimetry can be used for the calculation of resting metabolic rate and the thermic effect of food. This sum can be subtracted from the total energy expenditure resulting from doubly labeled water to directly measure energy expenditure associated with PA.

***Motion sensors***

***Pedometers:*** Pedometers are motion sensors attached to the waist which are triggered during walking. Pedometers can measure the total number of steps taken, but do not assess intensity or pace (5). However, research has established that walking approximately 100 steps per minute translates to a moderately intense level of PA (26). Applying this criterion, adults would need to achieve 3,000 steps in 30 minutes to satisfy the daily recommend quantity for medium PA. In addition, Tudor-Locke & Bassett (24) have also established pedometer cut-points for daily recommended PA. This concept extends beyond the classifications delineated in the PAGA, by including all PA accumulated in a given day. In this light, they recommend the following cut-points for adults free of disabilities or chronic disease:

- Sedentary: less than 5,000 steps per day
- Low active: Between 5,000–7,499 steps per day
- Somewhat active: Between 7,500–9,999 steps per day
- Active: Between 10,000 and 12,499 steps per day
- Highly active: Anything greater than or equal to 12,500 steps per day

More research is required to determine appropriate pedometer-driven step ranges across diverse populations. A review of 23 cross-sectional studies identified the following recommendations for specific populations:

- 8- to 10-year-old boys and girls: 12,000 to 16,000 steps per day
- Younger adults: 7,000 to 13,000 steps per day
- Older adults that are healthy: 6,000 to 8,500 steps per day
- People with disabilities and chronic illnesses: 3,500 to 5,500 steps a day

Pedometer cut points appear to have a direct effect on motivation to increase PA. Pedometers provide participants with a tangible goal to strive towards. As a self-monitoring tool, pedometers also appear to increase PA self-efficacy by offering immediate feedback (25). Bravata et al. (7) conducted a review of pedometer-based randomized control trials and found that pedometers were effective in increasing PA levels by 2,200 to 2,500 steps per day. They also found that pedometers were associated with significant decreases in body mass index and blood pressure. From a social marketing perspective, a number of programs have promoted 10,000 steps per day as an intervention outcome measure with favorable results (20). The long-term efficacy of pedometers to increase PA beyond intervention has not been substantiated (20).

Pedometers are popular for direct measurement, primarily because of their low cost and ease of use; however, the accuracy of pedometer measurement varies greatly by model and brand (20). During the course of an intervention, participants may lose or break pedometers. Consequently, the purchase of additional units beyond the minimal requirement to conduct the intervention should be factored into the research budget. Pedometers do not store data for later retrieval. Interventions that require participants to read and record their number of daily steps are at an increased risk for self-report biases. An additional potential drawback is that pedometers may be less accurate when attempting to quantify steps by individuals with altered gait patterns (7).

A standard protocol for measuring PA with a pedometer is to have participants wear the devices for 5 to 7 days and calculate the mean number of steps per day (20). This approach reduces measurement error by accounting for variability in PA between weekdays and weekends. At a minimum, researchers should aim to have participants wear the pedometer for at least three consecutive days. This time span for data collection has been shown to result in intraclass correlation coefficient reliability values of 0.80 for estimating steps per day (24).

***Accelerometers:*** Accelerometers are electronic motion sensors that measure body movement (5). These devices use an electronic component to assess acceleration of the body in a specific dimension. Accelerometers can be classified as uniaxial, biaxial, or triaxial, depending on the number of dimensions in which body movement can be monitored (27). Uniaxial monitors record vertical acceleration in one plane, biaxial monitors record acceleration in two directions (vertical and mediolateral), and triaxial monitors record acceleration in three directions (vertical, mediolateral, and anterior-posterior). Ideally, accelerometers should be small, light-weight, unobtrusive, sensitive, and able to store data over long periods of time (24).

Accelerometer data is expressed as counts per unit of time (24). Accelerometer counts are a measure of frequency and intensity of accelerations and decelerations. Activity counts can be translated into energy expenditure and PA intensity. Published cut-points for free-living PA are highly variable. The energy expenditure prediction equations used to estimate intensity levels vary depending on the calibration activities performed and the settings of these activities (7). An additional limitation of accelerometers is that high-quality brands are expensive and require training to interpret the data. In general, accelerometers tend to underestimate energy expended during inclined walking, running, swimming, rowing, cycling, and upper body exercises (27). Studies have found that children are not always compliant with accelerometer prescriptions due to stigmatization, as well as the obtrusiveness associated with the elastic belts designed to keep the monitor fixed into place. One advantage of accelerometers is that the units can store data for retrieval at a later time. Also, accelerometer data is not readily accessible to participants, which prevents the self-report bias that may accompany pedometers. Many models have wireless capabilities and can transmit data while the participant is wearing the device (6).

Research has shown that 4 to 12 measurements are required to gauge reliable accelerometry estimates of habitual PA in adults (6). Matthews found that 7 days of monitoring was required for estimating PA patterns of adults (18). Trost (22) found 8 to 9 days was required in adolescents and 4 to 5 days in children in order to achieve intraclass correlation reliability level of 0.80. The number of days for monitoring should be based on the battery life of the unit and the capacity of the unit for storing information (20).

***Global positioning system and global geographic information systems***

Environmental monitoring devices are gaining in popularity. Emerging technologies such as global positioning systems (GPS) and geographic information systems (GIS) are being integrated into accelerometer-based research to better understand the built PA environment (19). GIS uses a computer system to store information about a given location and the surrounding environment. GIS can spatially map and analyze data to examine the relationship between PA and neighborhood environmental factors such as land-use, walkability, recreational facilities, and parks. GPS uses a constellation of satellites to calculate geographic location and to track a specific activity. GPS can provide data on altitude, distance, time, and velocity. These two direct monitoring techniques are often used in conjunction with each other to provide a complete picture of environmental factors related to PA. Disadvantage of GPS devices include that they often fail to capture data related to indoor activity, and that they have the potential to fail under heavy tree canopy and dense urban areas (19).

***Heart rate monitors:*** Heart rate is the most convenient physiological marker for evaluating PA in the field (22). The heart rate monitor (HRM) is a small and unobtrusive instrument that measures electrical activity of the heart via a chest strap transmitter. The transmitter sends electrocardiograph signals to a digital receiver worn on the wrist. The majority of heart rate monitors can record and collect data at specified intervals, providing a description of intensity, frequency, and duration of PA.

Heart rate has a linear relationship with oxygen uptake, which allows for the estimation of an individual's energy expenditure, called the flex heart rate point (4). However, the oxygen consumption relationship must be determined for each individual in the study for several exercises. Even with customization and complex modeling, heart rate monitoring can be highly variable, and it frequently produces inaccurate estimates of energy expenditure when validated against doubly labeled water (4). There is also a lack of consensus on defining the flex heart rate point. To avoid problems at low intensity levels, monitoring should only be used to assess time spent in moderate and vigorous activity (27).

One advantage of the heart rate monitor is that the unit of measurement, heartbeat, can be used as a motivational tool in interventions. Similar to pedometers and daily steps, heartbeat can provide a tangible goal for participants to strive for (4). The major disadvantage of HRMs is that the heart rate can be affected by factors independent of PA, such as emotional stress, gender, training status, high ambient temperature, high humidity, and hydration. Additionally, data from HRMs can be lost due to signal interruptions and delayed heart rate responses.

A relatively novel method for PA assessment is the combined an accelerometer and heart rate monitor (11). When combined, the unique features of each device negate some of the disadvantages inherent in each device alone; furthermore, the measurement errors from the two devices are uncorrelated (11). At lower levels of intensity, the HRM is less accurate at estimating energy expenditure; whereas, it is during these periods when accelerometers are most accurate. Collectively, these two measurement tools have been shown to increase the accuracy of PA energy expenditure predictions by up to 20% (6).

***Direct observation:*** Behavioral observation is a method of measurement in which a trained observer classifies PA by recording activity for a pre-determined duration of time (17). Observations occur in the participants' natural settings and recordings are entered into either a standardized computer-based or paper-and-pencil entry form. Direct, or third-party, observation offers several advantages over other direct methods. Primarily, it can provide qualitative data related to environmental and psychosocial factors. It can also identify the type and intensity of PA. The main disadvantage of behavioral observation is the time-intensive requirement to conduct observation and to code the subsequent data. Also, the observer(s) must undergo training to ensure accurate recording.

Third-party observation is especially popular for measuring children's PA. Children under 10 years of age are unable to accurately or reliably report their own PA behaviors (17). Also, unlike adults, children's PA typically occurs in bouts of short duration which makes it difficult for older children to recall their PA behaviors. Proxy-report, such as having a parent or teacher report the child's PA behavior, is one alternative; however, teachers may be overburdened and parents may not be aware of the PA their child engaged in during the school day. Given these barriers, direct observation is not an attractive method for assessing children's PA.

A number of observation systems are available including CARS (Children's Activity Rating Scale), OSRAC-P (Observational System for Recording Activity in Children—Preschool Version), SOPLAY (System for Observing Play and Leisure Activity in Youth), and SOFIT (System for Observing Fitness Instruction Time) to assess the PA behaviors of children (3, 17). Direct observation has been demonstrated to be a valid and reliable method for measuring PA. Furthermore, CARS has been validated against a variety of direct monitoring methods including indirect calorimetry, accelerometry, and heart rate monitoring (17). CARS, OSRAC-P, and SOPLAY have each demonstrated significant interobserver reliability coefficient values of greater than 0.80 (8).

***Indirect monitoring:*** Indirect monitoring measures PA through self-report through mediums such as questionnaires, interviews, and activity logs (1). Indirect methods for monitoring PA are frequently employed due to their practicality, low cost, low response burden, and acceptance (1). Outside of their economic benefits, the primary advantage of indirect measurement is the contextual nature of the data. As opposed to direct field methods, indirect monitoring can provide qualitative details regarding the type, context, and setting for PA. While indirect methods are more economically feasible than direct methods for assessing PA, however, indirect monitoring is also susceptible to recall and response biases which raises concerns about the reliability and validity of the measures (17).

***Activity logs.*** Activity logs are detailed accounts of PA behaviors and patterns (3). Dependent on the specific research questions, participants can self-report (3, 5):

- Type of PA, i.e., brisk walking, gardening, or house cleaning
- Purpose for the PA, i.e., house maintenance, transportation
- Duration of PA, i.e., typically in minutes
- Intensity of the PA, i.e., light, moderate, vigorous
- Body position during PA, i.e., sitting, standing, walking
- Domain of PA, i.e., leisure time PA, occupational PA
- Sedentary behaviors, i.e., screen time, time sitting at work

The participant would maintain the log book for the defined observational period specified by the researcher. The time frame typically ranges from several days to a week to provide an average frequency of the activities described. Seasonal records may also be employed to obtain information about habitual PA and how seasonal variations interplay with PA behaviors.

Dependent on the type of research question and instrument framework, the recorded behaviors can range from writing down each activity as it is completed (sometimes referred to as a PA log), to specifying the participant record activities at stipulated time intervals (sometimes referred to as a PA diary). Researchers can attempt to calculate total energy expenditure based on the PA behaviors recorded in the log by assigning standardized MET values to each activity (2). Energy expenditure can then be estimated applying the metabolic formula: Kilocalories = MET x hours of activity x bodyweight in kilograms (3). While activity logs provide in-depth analysis of PA behaviors, they place a high burden on participants and researchers. As such, activity logs are generally better suited for tertiary prevention or for specific health conditions where detailed information is required for clinical purposes (3).

***Self-report questionnaires:*** Self-report questionnaires are the most frequently applied PA measurement device (5). Self-report assessment can be delivered through a variety of methods including face-to-face interviews, telephone interviews, mail-back questionnaires, hand-delivered questionnaires, and web-based questionnaires. In terms of PA assessment, there are three broad categories of questionnaires: global questionnaires, quantitative activity histories, and recall questionnaires.

Global questionnaires, such as those used as part of the Behavioral Risk Factor Surveillance State Questionnaire (10) are brief surveys that provide information related to general PA levels. This form of questionnaire is designed to stratify a population into general classifications; for example, active or inactive (3). Global questionnaires generally only tap into a few domains of PA. Global questionnaires can range from 1 to 4 items in length and only take a minute or so to complete. Global questionnaires do not provide details regarding specific PA behavior such as type and frequency (3). Instead, they are more suited to surveillance and tend to inquire about PA over a longer period of time, reducing systematic biases related to day of the week or season of the year.

Quantitative activity histories are typically the longest PA assessment questionnaires and can range from 15 to 80 items in length. A frequently used quantitative history questionnaire is the Minnesota Leisure-Time Physical Activity Questionnaire (21). Quantitative histories are detailed in nature and capture PA data related to the frequency and duration of activities performed over the past year or lifetime. Activity histories are useful for analyzing relationships between PA patterns and chronic disease outcomes.

Recall questionnaires are the most frequently applied format for PA assessment (3). Popular recall questionnaires applied in health promotion and education research include the International Physical Activity Questionnaire (12) and the Physical Activity Questionnaire for Older Children (16). Recall questionnaires can be used to identify details about the frequency, duration, and types of PA performed. Energy expenditure can also be estimated using standardized MET values (2). Recall time can vary from 24 hours to one year. The outcome measures for recall questionnaires can include ordinal scales to assess levels of PA or continuous data to determine whether public health activity guidelines are met within a given population. Recall self-report questionnaires are popular in health education research because they are relatively easy to administer, inexpensive, and generally well received by participants (1). Notwithstanding, there are several limitations with recall self-report.

Self-report creates the opportunity for response bias (5). Social desirability bias occurs when a participant responds to a questionnaire in a fashion that they interpret to be socially desirable, or that will paint them in a positive light. To minimize this bias, efforts should be made to ensure that participants understand that their responses are confidential and/or anonymous. Web-based delivery can help to increase the perception of confidentiality. If the instrument is administered in a closed setting, such as a classroom, a neutral third-party should dispense and collect the questionnaires to minimize situational contaminants. Additionally, social desirability scales, such as the Marlowe-Crowne Social Desirability Scale, can be integrated into the instrument to detect potentially invalid responses (13).

The cognitive demands of the recall process itself can also lead to biases (15). Roughly, the length of recall is inversely proportional to the level of recall precision. Recall is especially problematic in children and in the elderly. Children’s PA occurs sporadically and in bouts, which makes it difficult for them to accurately recall their PA behavior (17). Proxy recall questionnaires, such as having a parent or teacher report a child’s PA behavior, are often employed to overcome the cognitive immaturity of children. The elderly may have cognitive dysfunctions which can make recall difficult (15). Furthermore, they may have health issues such as poor eyesight which can make completing a hardcopy or web-based questionnaire especially challenging. Given the variable nature of the elderly population, researchers should investigate the particular segment they are interested in studying for any potential cognitive or physical limitations. Based on specific limitations of the sample, the researcher should consider face-to-face interviews for data collection.

To provide accurate measurement, PA recall questionnaires should be valid and reliable (5). Reliability is the stability of a measure. A reliable instrument should produce similar results over repeated administrations. In other words, sampling error should not impact the PA measurements (5). Validity is the extent to which an instrument measures what it intends to measure; in this case, PA. Ideally, PA recall questionnaires should undergo criterion validity in which they are tested against an objective method which provides the same outcome measure (5). For example, if the researcher wants to examine total minutes of moderately intense PA, one method would be to compare self-reported minutes against accelerometry. Or, if the researcher is interested in energy expenditure, they could validate the questionnaire against doubly labeled water or indirect calorimetry. It is also important to recognize that recall questionnaires are demographically dependent. Consequently, recall self-reports validated in one ethnic group are not necessarily valid in a different ethnic group (3).

## RECOMMENDATIONS

The accurate and reliable measurement of PA is central to health education and promotion research and evaluation practice. From a research perspective, PA measurement allows researchers to understand the correlates and determinants of PA (5). In terms of evaluation, measurement allows researchers to gauge the efficacy and effectiveness of interventions designed to increase or sustain PA. Epidemiological measurement of PA at a population level is beneficial for tracking the effects of health promotion and education interventions over time. From a health promotion perspective, accurate PA measurement provides evidence for initiatives in health promotion policy and practice.

The challenge of PA measurement is to apply assessment techniques that will provide accurate and reliable data relevant to the given program. Several direct and indirect methods for monitoring PA are available to researchers. PA monitoring techniques exist along a continuum of precision, cost, sophistication, and administrative/logistic burden (3). Generally, the cost of an assessment method is inversely proportional to its accuracy (18). Ideally, the applied method should be suitable to measure PA over long enough periods of time to be representative of normal daily life in free-living conditions, with minimal discomfort and interference to the participants, and applicable to large sample sizes (7). The ultimate choice of measurement modality is often a matter of optimization based on the study's research question(s), feasibility, and availability of resources (3).

## REFERENCES


1. Adamo KB, Prince SA, Tricco AC, Connor-Gorber S, and Tremblay M. A comparison of indirect versus direct measures for assessing physical activity in the pediatric population: a systematic review. *International Journal of Pediatric Obesity* 4: 2-27, 2009. Ainsworth BE. Compendium of physical activities: an update of activity codes and MET intensities. *Medicine and science in sports and exercise* 32: S498, 2000.
2. Ainsworth BE. How do I measure physical activity in my patients? questionnaires and objective methods. *British Journal of Sports Medicine* 43: 6-9, 2009.
3. Barreira TV, Kang M, Caputo JL, Farley RS, and Renfrow MS. Validation of the Actiheart monitor for the measurement of physical activity. *Int J Exerc Sci* 2: 60-71, 2009.
4. Bauman A, Phongsavan P, Schoeppe S, and Owen N. Physical activity measurement-a primer for health promotion. *Promotion & Education* 13: 92-103, 2006.
5. Berlin JE, Storti KL, and Brach JS. Using activity monitors to measure physical activity in free-living conditions. *Physical Therapy* 86: 1137-1145, 2006.
6. Bravata Dm S-SCSV and et al. Using pedometers to increase physical activity and improve health: A systematic review. *JAMA: The Journal of the American Medical Association* 298: 2296-2304, 2007.
7. Brown WH, Pfeiffer KA, McIver KL, Dowda M, Almeida MJCA, and Pate RR. Assessing preschool children's physical activity: the Observational System for Recording Physical Activity in Children-Preschool Version. *Research quarterly for exercise and sport* 77: 167-176, 2006.
8. Caspersen CJ, Powell KE, and Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Reports* 100: 126-131, 1985.
9. Centers for Disease Control and Prevention (CDC). *Behavioral Risk Factor Surveillance System Survey Questionnaire*. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2013.
10. Corder K, Brage Sr, Wareham NJ, and Ekelund U. Comparison of PAEE from combined and separate heart rate and movement models in children. *Medicine and science in sports and exercise* 37: 1761-1767, 2005.
11. Craig CL, Marshall AL, Sjorstrom M, et al. International physical activity questionnaire: 12-country reliability and validity. *Medicine and science in sports and exercise* 2003;35(8):1381-1395.
12. Crowne DP and Marlowe D. A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology* 24: 349-354, 1960.
13. DeLany JP. Measurement of energy expenditure. *Pediatric Blood & Cancer* 58: 129-134, 2012.
14. Forsén L, Loland NW, Vuillemin A, Chinapaw MJM, van Poppel MNM, Mokkink LB, van Mechelen W, and Terwee CB. Self-administered physical activity questionnaires for the elderly: a systematic review of measurement properties. *Sports Medicine* 40: 601-623, 2010.
15. Kowalski KC, Crocker PR, Faulkner RA. Validation of the Physical Activity Questionnaire for Older Children. *Pediatric exercise science* 1997;9:174-186.
16. Loprinzi PD and Cardinal BJ. Measuring children's physical activity and sedentary behaviors. *Journal of Exercise Science & Fitness* 9: 15-23, 2011.
17. Matthews C. *Use of self-report instruments to assess physical activity*. Champaign, Ill: Human Kinetics Inc, 2002.
18. Rodríguez DA, Brown AL, and Troped PJ. Portable global positioning units to complement accelerometry-based physical activity monitors. *Medicine and science in sports and exercise* 37: S572-581, 2005.
19. Schneider PL, Bassett JDR, Thompson DL, Pronk NP, and Bielak KM. Effects of a 10,000 steps per day goal in overweight adults. *American Journal of Health Promotion* 21: 85-89, 2006.
20. Taylor HL, Jacobs Jr DR, Schucker B, Knudsen J, Leon AS, Debacker G. A questionnaire for the assessment of leisure time physical activities. *Journal of chronic diseases* 1978;31(12):741-755.
21. Trost SG. Objective measurement of physical activity in youth: current issues, future directions. *Exercise and Sport Sciences Reviews* 29: 32-36, 2001.

22. Tudor-Locke C and Bassett D. How many steps/day are enough?: preliminary pedometer indices for public health. *Sports Medicine* 34: 1-8, 2004.
23. Tudor-Locke C, Sisson SB, Collova T, Lee SM, and Swan PD. Pedometer-determined step count guidelines for classifying walking intensity in a young ostensibly healthy population. *Canadian Journal of Applied Physiology* 30: 666-676, 2005.
24. Tudor-Locke CE and Myers AM. Challenges and opportunities for measuring physical activity in sedentary adults. *Sports Medicine* 31: 91-100, 2001.
25. US Department of Health and Human Services; 2008. Physical Activity Guidelines for Americans. [Accessed, September 8, 2013]. http://www.health.gov/paguidelines/
26. Warren JM, Ekelund U, Besson H, Mezzani A, Geladas N, and Vanhees L. Assessment of physical activity-a review of methodologies with reference to epidemiological research: a report of the exercise physiology section of the European Association of Cardiovascular Prevention and Rehabilitation. *European Journal of Cardiovascular Prevention & Rehabilitation* 17: 127-139, 2010.