Published in final edited form as:

J Sports Sci. 2017 February ; 35(3): 255–261. doi:10.1080/02640414.2016.1161218.

The accuracy of the 24-hour activity recall method for assessing sedentary behavior: the Physical Activity Measurement Survey (PAMS) project

Youngwon Kim^{1,2} and Gregory J. Welk²

¹MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Box 285 Institute of Metabolic Science, Cambridge Biomedical Campus, Cambridge, Cambridgeshire, UK CB2 0QQ

²Department of Kinesiology, Iowa State University, 257 Forker Building, Ames, IA, USA 50010

Abstract

Background—Sedentary behavior (SB) has emerged as a modifiable risk factor, but little is known about measurement errors of SB. The purpose of this study was to determine the validity of 24-hour Physical Activity Recall (24PAR) relative to SenseWear Armband (SWA) for assessing SB.

Methods—Each participant (n=1485) undertook a series of data collection procedures on two randomly selected days: wearing a SWA for full 24-hours, and then completing the telephone-administered 24PAR the following day to recall the past 24-hours' activities. Estimates of total sedentary time (TST) were computed without the inclusion of reported or recorded sleep time. Equivalence testing was used to compare estimates of TST.

Results—Analyses from equivalence testing showed no significant equivalence of 24PAR for TST (90% CI: 443.0 and 457.6 min/day) relative to SWA (equivalence zone: 580.7 and 709.8 min/day). Bland-Altman plots indicated individuals that were extremely or minimally sedentary provided relatively comparable sedentary time between 24PAR and SWA. Overweight/obese and/or older individuals were more likely to under-estimate sedentary time than normal weight and/or younger individuals.

Conclusion—Measurement errors of 24PAR varied by the level of sedentary time and demographic indicators. This evidence informs future work to develop measurement error models to correct for errors of self-reports.

Keywords

sedentary behavior; validity; physical activity recall; accelerometer; adults

Introduction

Sedentary behavior (SB) has emerged as a modifiable risk factor for mortality and cardiometabolic diseases, independent of accumulation of physical activity (PA) (Owen, Sparling, Healy, Dunstan, & Matthews, 2010). Evidence indicates that contemporary adults are known to spend more than half of their waking hours being sedentary (Matthews et al., 2008). However, different patterns of disparities in sedentary time among various adult populations are likely to be observed depending upon the choice of measurement tools (subjective *versus* objective) (Healy et al., 2011). This is mainly attributable to the different magnitude and distributions of measurement errors associated with the subjective and objective measurement tools. Research in this area is further compounded by the different ways to conceptualize SB and by differential influences of social/demographic characteristics on objective and subjective tools. To advance research on SB, therefore, it is important to better understand the underlying nature and magnitude of measurement errors, and their interplay with different socio-demographic indicators.

Efforts have been made to improve the accuracy and utility of measurement tools for assessing SB, but a key gap in SB research is the lack of consensus on the definition of SB (Gibbs, Hergenroeder, Katzmarzyk, Lee, & Jakicic, 2015). This makes it particularly challenging to properly assess SB and to compare prevalence and effects of SB across different studies. For example, a seminal paper (Pate, O'Neill, & Lobelo, 2008) originally defined SB as any activities inducing energy expenditure (EE) not substantially higher than resting EE (usually defined as <1.5METs). However, this definition does not take into account the behavioral posture of SB, and can potentially include time spent sleeping. Some studies have examined SB in terms of the postural characteristics of SB, such as time spent sitting (Katzmarzyk, Church, Craig, & Bouchard, 2009), while other studies have incorporated behavioral indicators such as screen time (Duncan, Vandelanotte, Caperchione, Hanley, & Mummery, 2012), TV viewing (Dunstan et al., 2005), and/or computer use (Ford, 2012). Concerted efforts have been made in recent years to establish consensus definitions of SB to help standardize research. The Sedentary Behaviour Research Network, for example, has specifically proposed a definition of SB "as any waking behavior characterized by an energy expenditure 1.5 METs while in a sitting or reclining posture" (Barnes et al., 2012). This definition specifically excludes sleep time in order to more directly target sitting time as the predominant component. Hence, it is common that SB research (Healy et al., 2007) uses overall classified minutes of sedentary time, which excludes sleep time. However, the quantification of sedentary time varies depending upon different measurement methods used.

Currently, a wide variety of objective and subjective measurement tools are available to assess SB. For instance, objective monitors (i.e. accelerometers, inclinometers) can provide reasonably accurate estimates of METs but most do not provide direct estimates of sitting and sleep. Subjective measurement tools (i.e. recalls, questionnaires) can provide contextual information about behavior but are more susceptible to measurement errors due to recall bias (Durante & Ainsworth, 1996), social desirability and social approval bias (Adams et al., 2005). Previous research (Atkin et al., 2012; Clark et al., 2009) has specifically documented the inherent limitations of longer-term recall instruments but short-term recall methods, such

as 24-hour Physical Activity Recall (24PAR), have shown promise for minimizing measurement errors in assessing PA and SB (Matthews, Moore, George, Sampson, & Bowles, 2012; Sallis & Saelens, 2000). Because recall of the past 24-hour period is fairly robust, the 24PAR is able to capture detailed behavioral patterns of SB, the 24PAR has promise for advancing research on SB. However, estimates must be compared against objective measures to evaluate validity. The Physical Activity Measurement Survey (PAMS) is a population-level study designed to understand measurement errors in self-report measures of PA but the design and measures also offer considerable potential to advance research on SB. Participants wore a SenseWear Armband (SWA) monitor for a 24-hour period and completed a detailed recall of behavior using an interviewer-administered 24PAR protocol. The use of both the subjective and objective measures is consistent with best practices in epidemiology research (Healy et al., 2011; Troiano, Pettee Gabriel, Welk, Owen, & Sternfeld, 2012) since it provides a more comprehensive evaluation of behavior. A previous study (Welk et al., 2014) using the PAMS data validated the 24PAR method against the SWA for PA but it is important to also examine the accuracy of the 24PAR for estimating SB. Therefore, the purpose of this study was to investigate the validity of the 24PAR tool relative to the SWA for assessing SB with the representative sample of adults from the PAMS project.

Method

Sample

The PAMS study used a multi-level stratified sampling strategy to obtain a representative sample of Iowa adults ages 20–75years. Potential participants were contacted through random digit dialing with a sample pool purchased from Survey Sampling International. The inclusion criteria were adults aged 20–75years, able to walk, and able to complete both a telephone and written paper survey in either English or Spanish. Any participants with medical conditions (i.e., allergies to the metal part of the armband monitor) preventing them from participation were excluded. The PAMS project was approved by the local Institutional Review Board at Iowa State University. All participants signed written informed consent before participation.

Instruments

24 Physical Activity Recall (24PAR): The 24PAR is a telephone- and interviewer-administered self-report tool designed to provide detailed information about activity time, energy expenditure, and specific context of SB and PA performed in the past day. With the 24PAR, participants are asked to report on activities of the past day from midnight to midnight in minimum bouts of 5 minutes across four distinct 6-hour time windows (early morning, morning, afternoon and evening). Each reported activity was matched with a corresponding MET score from the refined Compendium of PA (Ainsworth et al., 2011).

SenseWear Armband Mini (SWA): The SWA (BodyMedia, Inc., Pittsburgh, PA) is a non-invasive, multi-sensor activity monitor designed to be worn on the triceps of an upper arm. It relies on various physiological sensors (heat flux, galvanic skin response, skin temperature, near body temperature) as well as a tri-axial accelerometer. Resulting parameters (i.e.

activity time, MET, speed/distance, etc.) are produced using a manufacturer's proprietary algorithm based on the composite values of the sensors and tri-axial accelerometer. Previous validation studies (Fruin & Rankin, 2004; Machac, Prochazka, Radvansky, & Slaby, 2013) have demonstrated the high accuracy of the SWA for estimating energy expenditure and activity time for adults. Moreover, a critical strength of the SWA is the ability of the SWA to correctly identify non-wear time (vs. wear time), which makes it particularly useful for field-based surveillance and/or intervention research. Given these strengths, the SWA served as a criterion method to determine the validity of the 24PAR herein. The latest version of Software v8.0 (coupled with algorithms v5.2) was used to process the SWA data.

Data Collection

Each participant was asked to undertake two separate sessions of the measurement protocol (referred to as Trial 1 and Trial 2 hereafter). Each measurement protocol was composed of wearing the SWA for a full 24 hours on a randomly selected day and completing a 24PAR the next day recalling activities performed in the past day (same as the day for SWA monitoring). Staff members visited each participant's house one day prior to the monitoring day in order to distribute a SWA monitor and provide instructions on the appropriate use of it. They measured participants' height and weight during the visits. Staff members made a follow-up visit (i.e. one day after the monitoring day) to the participant's house to collect and download the SWA as well as schedule the second 24PAR/SWA session (i.e. Trial 2). A team of trained interviewers utilized a computer-assisted telephone interview system programmed using the Blaise Trigram methodology to carry out 24PAR interviews over the telephone. Each participant was asked to report on the types, minutes, purpose and location of each activity that they performed in the past day, the same day as the SWA monitoring day. The interviewers tracked total reported minutes to ensure that the sum of the reported minutes is 1,440 (i.e. 60min×24h). Participants reported on their socio-demographic information during the 24PAR interview. It took an average of 20min (range from 12 to 45min) to complete a single 24PAR interview. The data were collected across 8 3-month quarters over a 2-year time span. More detailed descriptions of the study design and data collection procedure are described elsewhere (Welk et al., 2014).

Data Processing

Minutes reported from the 24PAR were aggregated for each reported activity for each participant. Corresponding MET scores (assigned from the reduced set of PA Compendium) were used to categorize each reported activity as sedentary (MET ≤ 1.5) versus non-sedentary activity (MET > 1.5). All minutes with MET ≤ 1.5 were summed up for each participant to provide an overall estimate of Total Sedentary Time (TST) plus sleep. Time spent sleeping was obtained from the reported minutes for an activity code called 'Sleep and nap'. TST was thus calculated by subtracting sleep time from 'TST plus sleep'. Parallel calculations were conducted with the SWA. The SWA produces MET values for every minute so time spent with METs ≤ 1.5 were summed to estimate TST plus sleep. Sleep time for the SWA was directly obtained from the Bodymedia software and was then subtracted from 'TST plus sleep' in similar ways to produce the corresponding estimates of TST. Epidemiological studies typically use self-reported sleep time when deriving "TST" from objective methods (i.e. the sum of objectively estimated 'TST+sleep' – self-reported sleep time). Therefore, we

additionally used this standard methodology to calculate another type of TST for SWA by subtracting self-reported sleep time (obtained from 24PAR) from 'TST+sleep' (obtained from SWA). The values from the two trials were averaged for 24PAR and SWA.

Statistical Analyses

The study examined factors influencing agreement between subjective and objective estimates of SB. The analyses adopted a similar equivalence testing methodology used in a previous PAMS paper to evaluate agreement on estimates of EE and PA (Welk et al., 2014). In this case, we used a weighted 95% equivalence test and weighted correlation analyses to evaluate group-level agreement between the estimates of SB from the 24PAR and SWA. In regards to the equivalence testing, if a 90% confidence interval (CI) for the 24PAR is completely included within a specified equivalence zone (i.e. $\pm 10\%$ of the mean of the SWA), the 24PAR is considered to be significantly equivalent to the SWA, on average. Greater detail about the equivalence testing method is provided elsewhere (Kim et al., 2014; Welk et al., 2014). Individual-level agreement was evaluated using weighted mean absolute percent errors (MAPE) and Bland-Altman plots with 95% limits of agreement.

Weighted Analyses of Variance (ANOVA) were performed to examine the interaction effects among age, BMI and gender for differences in TST between the 24PAR and SWA. Variables and interactions with significant beta coefficients (i.e. p -values < 0.05) were included in the model. The resulting interaction effects were illustrated on heat maps.

All analyses were adjusted to account for the sample weights. The survey weights were calculated with a customized jackknife procedure to estimate standard errors (Fuller, 2009), and used to account for the complex sampling design of the PAMS project, and to obtain population-level estimates. Data processing and management were performed in STATA/SE version 12 (StataCorp LP, College Station, TX). Equivalence testing and heat map analyses were performed in SAS (SAS Institute Inc, Cary, NC) and R, respectively.

Results

A final sample of 1,458 was included for analyses in this study. Of the 1,458 adults, 1,356 provided the 24PAR and SWA data for both Trial 1 and 2, and 102 provided the data only for one of the two trials (i.e. 71 with only Trial 1 and 31 with only Trial 2). Specific characteristics of the final sample of the participants are presented in Table 1.

Results from 95% equivalence testing revealed that the 24PAR yielded substantial underestimation of TST (90% CI: 443.0 and 457.6 min/day) relative to the equivalence zone of the SWA (580.7 and 709.8 min/day), which led to insignificant equivalence. (See Figure 1; Top panel)

A significant moderate correlation coefficient was observed for TST ($R=0.45$). A relatively large MAPE value was identified for TST (35.1%). (See Table 2) In analyses to examine if the TST of the 24PAR is comparable to the TST of the SWA with self-reported sleep time, similar patterns of results were identified. For example, the 24PAR (90% CI: 443.0 and 457.6 min/day) was not equivalent to the SWA (equivalence zone: 508.0 and 621.5 min/day)

on average (See Figure 1; Bottom Panel), and the correlation and MAPE were 0.49 and 120%, respectively (See Table 2). Bland-Altman plot analyses revealed a pattern of systematic bias. To be specific, a “wide diamond” shape was observed, implying that individuals with extremely low or high levels of sedentary time provided comparable sedentary time estimates with the 24PAR in relation to the SWA. However, those that are moderately sedentary yielded large recall errors (primarily under-estimation) with the 24PAR. Corresponding 95% limits of agreement were -618.6 and 176.0 min/day for TST, which is deemed very wide. (See Figure 2)

Figure 3 illustrates 3-way interaction effects of age, gender and BMI for differences in TST on heat maps. White color indicates that the differences between the 24PAR and SWA are zero. Blue color indicates that values from the 24PAR are smaller than the SWA, and red color indicates that values from the 24PAR are larger than the SWA. The vast majority of the heat map areas were mostly depicted in blue, indicating under-estimation of sedentary time by the 24PAR. More specifically, individuals that are considered overweight/obese (i.e. BMI>25) and/or relatively older (i.e. ages>50years) were more likely to under-estimate TST in comparison with those that are normal weight and/or relatively younger. No gender effects were identified. (See Figure 3)

Discussion

Accurate assessment of SB is critical to advancing research on SB (Sallis, Owen, & Fotheringham, 2000). Key priorities include a more thorough characterization of the magnitude, sources and distributions of measurement errors in estimates of SB. The inherent challenges of assessing SB with self-report methods are due, in large part, to the inconsistent quantification of SB and the inability to control for individual variability in perceptions and reporting of SB in the population. In this regard, this study helps to fill this gap by capitalizing on a representative sample of adults to determine the patterns and distributions of measurement errors for the 24PAR relative to objective estimates from the SWA by age, gender and BMI.

There are no previous studies using this particular version of the 24PAR but insightful comparisons can be made with several past studies (Calabro et al., 2009; Matthews et al., 2013; Welk et al., 2014) that have adopted the basic 24PAR format. To be specific, Calabro et al. examined the validity of the computerized 24PAR tool in relation to SenseWear Pro2 Armband monitors and IDEEA (i.e. pattern recognition technology) in a convenient sample of 20 adults (Calabro et al., 2009). Overall, the 24PAR exhibited good agreement for estimating total EE (correlations ranging from 0.89–0.91) but moderate agreement for assessing moderate-to-vigorous physical activity (MVPA) time (correlations ranging from 0.48–0.70), relative to both the Pro2 Armband and IDEEA. Similar patterns of results were observed from Welk et al. in a previous study from the PAMS project (Welk et al., 2014). The 24PAR yielded relatively small MAPEs for EE (ranges from 10.3% to 15%) and large MAPEs for MVPA (ranges from 68.6% to 269.5%) compared to estimates from the SWA. The only comparable study evaluating estimates of SB is by Matthews et al. which assessed the accuracy of TST relative to activPAL in a convenient sample of 88 adults (Matthews et al., 2013). The agreement statistics from the current study (i.e. correlation of 0.45 and non-

equivalent estimates of TST) are similar to those from Matthews et al. study (i.e. correlations of 0.34 and 0.67 and significant differences of TST) (Matthews et al., 2013). Recent research by Gomersall et al. compared another type of 24-hour recall methods, called MARCA, against activPAL for assessment of sitting/lying time in a sample of 58 adults (Gomersall, Pavey, Clark, Jasman, & Brown, 2015). They found somewhat low agreement, with a correlation of 0.77 and wide limits of agreement for Bland-Altman plots. Additional research is clearly needed to further investigate the accuracy of the 24PAR specific for assessment of SB in relation to a true gold standard method and under various environments. The use of the PAMS data herein allowed for identifying the underlying patterns and distributions of measurement errors of the 24PAR in assessing SB at the population level. A notable pattern (detected from the Bland-Altman plots) is the tendency that individuals that were highly or minimally sedentary exhibited relatively smaller measurement errors with the 24PAR, in comparison with those that were moderately sedentary. This finding makes conceptual sense since it would be easier for an individual to correctly recall the previous day's sedentary activities if the total amount of time spent sedentary was extremely large or small. Another interesting finding was the variation of measurement errors by age, gender and BMI (identified from the heat map analysis). To be specific, individuals that were older than 50years old and/or overweight/obese were more likely to under-estimate sedentary time in comparison with those that were younger than 50years old and/or normal weight. This result clearly confirms the known concept of relative intensity (Howley, 2001). For example, Welk et al. using the same PAMS data indicated that older and/or obese individuals were more likely to perceive a low intensity of activities as moderate or vigorous (Welk et al., 2014). As a result, substantial overestimation of MVPA was observed in that study (Welk et al., 2014). Interestingly, the overestimation of MVPA by the older and/or obese individuals partly explains the underestimation of overall SB in the current study. The dependency of these reported estimates suggests the need to develop and calibrate separate measurement error models to correct for recall biases in the estimation of both MVPA and SB.

The current study is the first empirical research to validate a self-report tool for assessing different constructs of SB at the population level. The 24PAR and SWA provided substantially different estimates of TST, and this is mainly attributable to the difference between the 24PAR and SWA in estimates of sleep time. For example, the TST by the SWA and 24PAR was 654.3min/day and 450.3min/day, respectively, which was due, in large part, to the considerable difference for sleep time: 361.2min/day by SWA *versus* 477.6min/day by 24PAR. Sleep time for the SWA is based on physiological changes in the sensors while the 24PAR is a subjective report. It is logical for the objective report of sleep time to be lower than the reported since it is not possible to directly recall when sleep starts. The wide disparity in these values reveals a major source of error that needs to be addressed to advance research on SB. While no research has been conducted to investigate the accuracy of the 24PAR for sleep detection, three studies (O'Driscoll, Turton, Copland, Strauss, & Hamilton, 2013; Segura-Jimenez et al., 2015; Sharif & BaHammam, 2013) have examined the validity of the SWA for estimating sleep time. However, they (O'Driscoll et al., 2013; Segura-Jimenez et al., 2015; Sharif & BaHammam, 2013) provided mixed results and included clinical populations (i.e. obstructive sleep apnea, fibromyalgia). Additional research is clearly needed with both objective and subjective measures to separate out the

positive contributions of sleep time from other sources of SB. Most accelerometer-derived estimates are based solely on lack of movement rather than physiological changes detected from the SWA, so it is important to better understand the nature of errors in methods for estimating sleep time in order to obtain more accurate estimates of TST.

This study provides new insights about error in estimates of SB. To our knowledge, this is the first empirical investigation using a representative sample of adults to examine the accuracy of a self-report tool for assessing SB. Another strength is the use of a series of novel statistical procedures to test agreement between the 24PAR and SWA. Using the equivalence testing allowed us to determine if there is “evidence of equivalence” between the two methods, and the heat map analyses illustrated complicated patterns of measurement errors associated with the 24PAR.

However, this study has some limitations which should be noted. First of all, the SWA is not capable of classifying standing activities, while the 24PAR does. Thus, we examined TST, which does not incorporate standing, in order to facilitate direct comparisons between the 24PAR and SWA. Time spent standing would further distinguish types of light activity that have been shown to have health benefits (Katzmarzyk, 2014), but the SWA is not presently able to distinguish standing from sitting and/or walking. Moreover, the SWA has not been systematically validated against strong reference methods in terms of assessing SB and differentiating sitting from other behavioral constructs. However, several previous studies have found high validity of the SWA for estimating energy expenditure (Fruin & Rankin, 2004; Machac et al., 2013), which is a main parameter (i.e. MET 1.5) used herein to quantify SB. Another limitation is that the vast majority of the participants were White (88%) and nearly all had at least a high school education (97%). The stratified sampling technique allowed us to oversample the minority populations (11%) but consideration was not given to education. Iowa has among the lowest rates in the nation for adults without a high school diploma (~10%) but findings from the study may not generalize to other ethnicities (i.e. Black, Hispanic, etc.) or to those without high school education.

Conclusions

The present study examined the agreement between the 24PAR and SWA for evaluating SB and the influences of demographic indicators on the agreement. Overall, the 24PAR exhibited relatively low validity for TST estimates. The magnitude and distributions of measurement errors varied considerably by demographic indicators (particularly, age and BMI). Evidence obtained herein provides unique insights about the underlying nature of measurement errors for assessment of SB. This research can serve as a fundamental framework for future work to establish measurement error models for minimizing errors inherent in the 24PAR for estimating SB in adult populations.

References

- Adams SA, Matthews CE, Ebbeling CB, Moore CG, Cunningham JE, Fulton J, Hebert JR. The effect of social desirability and social approval on self-reports of physical activity. *Am J Epidemiol.* 2005; 161(4):389–398. doi: 161/4/389 [pii] 10.1093/aje/kwi054. [PubMed: 15692083]

- Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR Jr, Tudor-Locke C, Leon AS. 2011 Compendium of Physical Activities: a second update of codes and MET values. *Med Sci Sports Exerc.* 2011; 43(8):1575–1581. DOI: 10.1249/MSS.0b013e31821ece12 [PubMed: 21681120]
- Atkin AJ, Gorely T, Clemes SA, Yates T, Edwardson C, Brage S, Biddle SJH. Methods of Measurement in epidemiology: Sedentary Behaviour. *International Journal of Epidemiology.* 2012; 41(5):1460–1471. DOI: 10.1093/ije/Dys118 [PubMed: 23045206]
- Barnes J, Behrens TK, Benden ME, Biddle S, Bond D, Brassard P, Network, S.B.R. Letter to the Editor: Standardized use of the terms “sedentary” and “sedentary behaviours”. *Applied Physiology Nutrition and Metabolism-Physiologie Appliquee Nutrition Et Metabolisme.* 2012; 37(3):540–542. DOI: 10.1139/H2012-024
- Calabro MA, Welk GJ, Carriquiry AL, Nusser SM, Beyler NK, Mathews CE. Validation of a computerized 24-hour physical activity recall (24PAR) instrument with pattern-recognition activity monitors. *J Phys Act Health.* 2009; 6(2):211–220. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/19420399>. [PubMed: 19420399]
- Clark BK, Sugiyama T, Healy GN, Salmon J, Dunstan DW, Owen N. Validity and reliability of measures of television viewing time and other non-occupational sedentary behaviour of adults: a review. *Obes Rev.* 2009; 10(1):7–16. DOI: 10.1111/j.1467-789X.2008.00508.x [PubMed: 18631161]
- Duncan MJ, Vandelanotte C, Caperchione C, Hanley C, Mummery WK. Temporal trends in and relationships between screen time, physical activity, overweight and obesity. *BMC Public Health.* 2012; 12 doi:Artn 1060. doi: 10.1186/1471-2458-12-1060
- Dunstan DW, Salmon J, Owen N, Armstrong T, Zimmet PZ, Welborn TA, AusDiab Steering C. Associations of TV viewing and physical activity with the metabolic syndrome in Australian adults. *Diabetologia.* 2005; 48(11):2254–2261. DOI: 10.1007/s00125-005-1963-4 [PubMed: 16211373]
- Durante R, Ainsworth BE. The recall of physical activity: Using a cognitive model of the question answering process. *Med Sci Sports Exerc.* 1996; 28(10):1282–1291. DOI: 10.1097/00005768-199610000-00012 [PubMed: 8897386]
- Ford ES. Combined television viewing and computer use and mortality from all-causes and diseases of the circulatory system among adults in the United States. *BMC Public Health.* 2012; 12 doi:Artn 70. doi: 10.1186/1471-2458-12-70
- Fruin ML, Rankin JW. Validity of a multi-sensor armband in estimating rest and exercise energy expenditure. *Med Sci Sports Exerc.* 2004; 36(6):1063–1069. doi:00005768-200406000-00021 [pii]. [PubMed: 15179178]
- Fuller, WA. *Sampling Statistics.* New York: Wiley; 2009.
- Gibbs BB, Hergenroeder AL, Katzmarzyk PT, Lee IM, Jakicic JM. Definition, measurement, and health risks associated with sedentary behavior. *Med Sci Sports Exerc.* 2015; 47(6):1295–1300. DOI: 10.1249/MSS.0000000000000517 [PubMed: 25222816]
- Gomersall SR, Pavey T, Clark BK, Jasman A, Brown WJ. Validity of a Self-report Recall Tool for Estimating Sedentary Behaviour in Adults. *J Phys Act Health.* 2015; doi: 10.1123/jpah.2014-0602
- Healy GN, Clark BK, Winkler EAH, Gardiner PA, Brown WJ, Matthews CE. Measurement of Adults' Sedentary Time in Population-Based Studies. *American Journal of Preventive Medicine.* 2011; 41(2):216–227. DOI: 10.1016/j.amepre.2011.05.005 [PubMed: 21767730]
- Healy GN, Dunstan DW, Salmon J, Cerin E, Shaw JE, Zimmet PZ, Owen N. Objectively measured light-intensity physical activity is independently associated with 2-h plasma glucose. *Diabetes Care.* 2007; 30(6):1384–1389. DOI: 10.2337/Dc07-0114 [PubMed: 17473059]
- Howley ET. Type of activity: resistance, aerobic and leisure versus occupational physical activity. *Med Sci Sports Exerc.* 2001; 33(6):S364–S369. DOI: 10.1097/00005768-200106001-00005 [PubMed: 11427761]
- Katzmarzyk PT. Standing and mortality in a prospective cohort of Canadian adults. *Med Sci Sports Exerc.* 2014; 46(5):940–946. DOI: 10.1249/MSS.0000000000000198 [PubMed: 24152707]
- Katzmarzyk PT, Church TS, Craig CL, Bouchard C. Sitting Time and Mortality from All Causes, Cardiovascular Disease, and Cancer. *Medicine and Science in Sports and Exercise.* 2009; 41(5):998–1005. DOI: 10.1249/Mss.0b013e3181930355 [PubMed: 19346988]

- Kim Y, Crouter SE, Lee JM, Dixon PM, Gaesser GA, Welk GJ. Comparisons of prediction equations for estimating energy expenditure in youth. *J Sci Med Sport*. 2014; doi: 10.1016/j.jsams.2014.10.002
- Machac S, Prochazka M, Radvansky J, Slaby K. Validation of physical activity monitors in individuals with diabetes: energy expenditure estimation by the multisensor SenseWear Armband Pro3 and the step counter Omron HJ-720 against indirect calorimetry during walking. *Diabetes Technology & Therapeutics*. 2013; 15(5):413–418. DOI: 10.1089/dia.2012.0235 [PubMed: 23557200]
- Matthews CE, Chen KY, Freedson PS, Buchowski MS, Beech BM, Pate RR, Troiano RP. Amount of time spent in sedentary behaviors in the United States, 2003–2004. *Am J Epidemiol*. 2008; 167(7): 875–881. DOI: 10.1093/aje/kwm390 [PubMed: 18303006]
- Matthews CE, Keadle SK, Sampson J, Lyden K, Bowles HR, Moore SC, Fowke JH. Validation of a previous-day recall measure of active and sedentary behaviors. *Med Sci Sports Exerc*. 2013; 45(8): 1629–1638. DOI: 10.1249/MSS.0b013e3182897690 [PubMed: 23863547]
- Matthews CE, Moore SC, George SM, Sampson J, Bowles HR. Improving self-reports of active and sedentary behaviors in large epidemiologic studies. *Exerc Sport Sci Rev*. 2012; 40(3):118–126. DOI: 10.1097/JES.0b013e31825b34a0 [PubMed: 22653275]
- O'Driscoll DM, Turton AR, Copland JM, Strauss BJ, Hamilton GS. Energy expenditure in obstructive sleep apnea: validation of a multiple physiological sensor for determination of sleep and wake. *Sleep Breath*. 2013; 17(1):139–146. DOI: 10.1007/s11325-012-0662-x [PubMed: 22318784]
- Owen N, Sparling PB, Healy GN, Dunstan DW, Matthews CE. Sedentary behavior: emerging evidence for a new health risk. *Mayo Clinic Proceedings*. 2010; 85(12):1138–1141. doi: 10.4065/mcp.2010.0444S0025-6196(11)60368-6[pil]. [PubMed: 21123641]
- Pate RR, O'Neill JR, Lobelo F. The evolving definition of “sedentary”. *Exerc Sport Sci Rev*. 2008; 36(4):173–178. doi:10.1097/JES.0b013e3181877d1a 00003677-200810000-00002 [pii]. [PubMed: 18815485]
- Sallis JF, Owen N, Fotheringham MJ. Behavioral epidemiology: a systematic framework to classify phases of research on health promotion and disease prevention. *Ann Behav Med*. 2000; 22(4):294–298. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11253440>. [PubMed: 11253440]
- Sallis JF, Saelens BE. Assessment of physical activity by self-report: status, limitations, and future directions. *Res Q Exerc Sport*. 2000; 71(2 Suppl):S1–14. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10925819>. [PubMed: 10925819]
- Segura-Jimenez V, Camiletti-Moiron D, Munguia-Izquierdo D, Alvarez-Gallardo IC, Ruiz JR, Ortega FB, Delgado-Fernandez M. Agreement between self-reported sleep patterns and actigraphy in fibromyalgia and healthy women. *Clinical and Experimental Rheumatology*. 2015; 33(1):S58–S67. Retrieved from <Go to ISI>://000352829200010.
- Sharif MM, BaHammam AS. Sleep estimation using BodyMedia's SenseWear (TM) armband in patients with obstructive sleep apnea. *Annals of Thoracic Medicine*. 2013; 8(1):53–57. DOI: 10.4103/1817-1737.105720 [PubMed: 23440703]
- Troiano RP, Pettee Gabriel KK, Welk GJ, Owen N, Sternfeld B. Reported physical activity and sedentary behavior: why do you ask? *J Phys Act Health*. 2012; 9(Suppl 1):S68–75. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/22287450>. [PubMed: 22287450]
- Welk GJ, Kim Y, Stanfill B, Osthus DA, Calabro MA, Nusser SM, Carriquiry A. Validity of 24-h physical activity recall: physical activity measurement survey. *Med Sci Sports Exerc*. 2014; 46(10):2014–2024. DOI: 10.1249/MSS.0000000000000314 [PubMed: 24561818]

Figure 1.

Results from 95% equivalence testing for agreement between the 24-hour Physical Activity Recall (24PAR) and SenseWear Armband (SWA) for Total Sedentary Time (TST). Note: Thick solid lines indicate 90% confidence intervals for the 24PAR, and double dotted lines indicate equivalence zones defined by the SWA. ^ indicates that TST for 24PAR was calculated as quantified sedentary time from 24PAR – sleep time from 24PAR, and that TST for SWA was calculated as quantified sedentary time from SWA – sleep time from SWA. * indicates that TST for 24PAR was calculated as quantified sedentary time from 24PAR – sleep time from 24PAR, and that TST for SWA was calculated as quantified sedentary time from SWA – sleep time from 24PAR.

Figure 2.
Bland-Altman plots with 95% limits of agreement for Total Sedentary Time (TST).

Figure 3.

Heat maps for 3-way interaction effects on differences in Total Sedentary Time between 24-hour physical activity recall (24PAR) and SenseWear Armband (SWA). Note: White area indicates no difference between the 24PAR and SWA. Blue and red area indicate under- and over-estimation of sedentary time, respectively, by the 24PAR relative to the SWA.

Table 1

Characteristics (mean, proportion, and standard error) of the participants.

Variables	All (n=1458)	Female (n=848)	Male (n=610)
<i>Age (years), Mean(SE)</i>	46.2 (0.4)	46.9 (0.6)	45.5 (0.6)
20–29 years	8.3 (1.0)	7.5 (1.5)	9.2 (1.5)
30–39 years	22.7 (1.7)	19.3 (2.2)	26.2 (2.5)
40–49 years	33.7 (1.8)	37.3 (2.6)	30.1 (2.5)
50–71 years	35.2 (1.5)	36.0 (2.1)	34.4 (2.3)
<i>Body Mass Index (BMI)</i>	30.0 (0.3)	29.6 (0.4)	30.3 (0.4)
Normal weight	24.7 (1.6)	30.4 (2.4)	18.9 (2.0)
Overweight	32.9 (1.7)	30.4 (2.3)	35.4 (2.5)
Obese	42.4 (1.8)	39.2 (2.4)	45.7 (2.6)
<i>Ethnicity</i>			
White	88.5 (1.3)	88.8 (0.2)	88.1 (1.9)
Black	7.1 (1.0)	8.0 (1.4)	6.1 (1.2)
Others	4.5 (1.0)	3.1 (1.1)	5.8 (1.6)
<i>Education</i>			
Less than a high school diploma	3.1 (0.6)	3.9 (1.1)	2.3 (0.6)
High school diploma/some college	50.2 (1.8)	50.0 (2.5)	50.4 (2.6)
College/graduate school	46.7 (1.8)	46.1 (2.5)	47.3 (2.6)
<i>Income</i>			
Less than \$25,000	13.7 (1.1)	15.6 (1.6)	11.9 (1.7)
From \$25,000 up to \$75,000	43.9 (1.8)	48.1 (2.6)	39.6 (2.5)
More than \$75,000	42.4 (1.9)	36.3 (2.6)	48.5 (2.7)
<i>Employment</i>			
Full-time job	66.5 (1.6)	54.8 (2.5)	78.2 (2.0)
Part-time job	12.2 (1.2)	16.9 (1.9)	7.4 (1.5)
Others	21.3 (1.3)	28.2 (2.1)	14.4 (1.6)

Note: All values were weighted to account for the complex sampling design. Values in parenthesis represent standard errors unless otherwise indicated. “Others” for Employment includes ‘Unemployed/Retired/Full-time homemaker or something else’

Abbreviation: SE – Standard Error

Table 2

Indicators for agreement between 24-hour physical activity recall (24PAR) and SenseWear Armband (SWA) monitor

Definitions of sedentary behavior	Least Squares Means (Minute)		Correlations	MAPE, %
	24PAR	SWA		
Total Sedentary Time (TST)	450.3 (4.4) ^a	654.3 (5.1) ^b	0.45 [*] (0.04)	35.1 (0.9)
Total Sedentary Time (TST)	450.3 (4.4) ^a	565.0 (5.5) ^c	0.49 [*] (0.04)	120.0 (35.8)

Notes: values in parentheses are standard errors unless otherwise specified.

Abbreviations: 24PAR – 24-hour Physical Activity Recall; SWA – SenseWear Armband; MAPE – Mean Absolute Percent Error

^{*} statistical significance (p-values less than 0.0001)

^a calculated as quantified sedentary time from 24PAR – sleep time from 24PAR

^b calculated as quantified sedentary time from SWA – sleep time from SWA

^c calculated as quantified sedentary time from SWA – sleep time from 24PAR

All values were estimated with the sampling weights applied.