Published in final edited form as:

Med Sci Sports Exerc. 2013 August ; 45(8): 1629–1638. doi:10.1249/MSS.0b013e3182897690.

Validation of a Previous-Day Recall Measure of Active and Sedentary Behaviors

Charles E. Matthews¹, Sarah Kozey Keadle², Joshua Sampson³, Kate Lyden², Heather R. Bowles⁴, Stephen C. Moore¹, Amanda Libertine², Patty S. Freedson², and Jay H. Fowke⁵

¹Nutritional Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD

²Department of Kinesiology, University of Massachusetts, Amherst, MA

³Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD

⁴Risk Factor Monitoring and Methods Branch, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD

⁵Division of Epidemiology, Department of Medicine, Vanderbilt University Medical School, Nashville, TN

Abstract

Purpose—A previous-day recall (PDR) may be a less error prone alternative to traditional questionnaire-based estimates of physical activity and sedentary behavior (e.g., past year), but validity of the method is not established. We evaluated the validity of an interviewer administered PDR in adolescents (12–17 years) and adults (18–71 years).

Methods—In a 7-day study, participants completed three PDRs, wore two activity monitors, and completed measures of social desirability and body mass index (BMI). PDR measures of active and sedentary time was contrasted against an accelerometer (ActiGraph) by comparing both to a valid reference measure (activPAL) using measurement error modeling and traditional validation approaches.

Results—Age- and gender-specific mixed models comparing PDR to activPAL indicated: (1) a strong linear relationship between measures for sedentary (regression slope = β_f =0.80 to 1.13) and active time (β_f =0.64 to 1.09); (2) person-specific bias was lower than random error; and (3) correlations were high (Sedentary: r = 0.60 to 0.81; Active: r = 0.52 to 0.80). Reporting errors were not associated with BMI or social desirability. Models comparing ActiGraph to activPAL indicated: (1) a weaker linear relationship between measures for sedentary (β_f =0.63 to 0.73) and active time (β_f =0.61 to 0.72); (2) person-specific bias was slightly larger than random error; and (3) correlations were high (Sedentary: r = 0.68 to 0.77; Active: r = 0.57 to 0.79).

Conclusions—Correlations between the PDR and activPAL were high, systematic reporting errors were low, and the validity of the PDR was comparable to the ActiGraph. PDRs may have

Corresponding Author: Charles E. Matthews, PhD, Nutritional Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Blvd, EPS 3028, Bethesda, MD 20892-7335, charles.matthews2@nih.gov, phone: 301.402.3491, fax: 301.496.6829.

Conflict of Interest

Patty S. Freedson is a member of the ActiGraph Scientific Advisory Board. No other potential conflicts of interest are declared. The results of the present study do not constitute endorsement by the American College of Sports Medicine.

value in studies of physical activity and health, particularly those interested in measuring the specific type, location, and purpose of activity-related behaviors.

Keywords

exposure assessment; measurement error; physical activity; behavioral epidemiology

INTRODUCTION

There has been extraordinary progress in the field of physical activity epidemiology in the last 50 years. Lack of participation in moderate-vigorous exercise (38), and more recently prolonged time spent in sedentary behavior—or sitting, have been associated with increased risk for mortality and chronic diseases (e.g., (30;45)), including certain cancers (e.g., (33)). Clearly, the exposure assessments employed in these studies, typically questionnaires that estimate usual amounts of physically active and sedentary behaviors (e.g., past year), have been successful in identifying many strong behavior-disease associations. On the other hand, these same questionnaires probably contain a substantial amount of measurement error (10;32), which leads to a loss of statistical power to test etiologic hypotheses, attenuation in the strength of the associations observed, and difficulties characterizing dose-response relationships that are critical in the development of evidence-based recommendations (42).

Better measurements are needed to address the limitations of traditional questionnaire-based tools, and recent summaries of the current state of the art of exposure assessment provide insight into why both device-based (12) and self-report methodologies (6) can and should play complementary roles in future studies. Recent systematic reviews have noted that few physical activity questionnaires have validity coefficients (correlations) greater than 0.5 (19;34;49) compared with objective measures, and sedentary behavior questionnaires have been noted to have a similarly modest level of validity (3;17).

Short-term recalls (e.g., diaries, previous day recalls) have been suggested as an alternative to traditional questionnaires that typically require longer term recall of behavior (32;35;48). New technologies such as web-based surveys and mobile devices coupled with emerging measurement error correction techniques (11;35) now make it feasible to use such methods as a primary exposure assessment tool in large scale studies. Previous day recalls (PDR) offer several advantages over questionnaire-based estimates of usual activity and sedentary behavior. First, they allow respondents to rely on episodic memory to generate reports about time spent in specific activity-related behaviors, rather than use of estimation strategies and long-term averaging (25). Thus, the information reported on the PDR may be more accurate. Second, PDRs capture more detailed information about different types of activities, offer a unique opportunity to assess body posture (i.e., sitting vs. standing), as well as information about behavioral context (e.g., location and purpose) not available from other measures. Hence, PDRs may be particularly valuable for studies interested in posture-based estimates of sedentary behavior, or that require information about where and why physically active and sedentary behaviors occur.

An important first step in establishing the proof of principle for PDRs for use in future studies is to test the validity of the method. Accordingly, the purpose of this report is to evaluate the validity of an interviewer administered PDR of physically active and sedentary behaviors in free-living adolescents and adults compared to the activPAL, an accurate and precise reference measure for distinguishing between active and sedentary behaviors (15;23). To provide insight into the measurement properties of the PDR compared to another instrument, we conducted a parallel analysis evaluating the validity of the ActiGraph

monitor compared to the activPAL. In secondary analyses, we also evaluated PDR measures of light and moderate-vigorous physical activity using common ActiGraph cut-points.

MATERIALS AND METHODS

Study Design

During the 7-day study period adolescents (12–17 years) and middle-aged adults (18–71 years) from Amherst, MA and Nashville, TN wore two activity monitors and received three unannounced telephone-administered PDRs (two weekdays, one weekend day). Eligible participants for the study were 12 to 75 years of age and were free of debilitating chronic diseases (e.g., heart failure, severe claudication, terminal cancer), major cognitive or psychiatric disorders (e.g., dementia, schizophrenia), and major orthopedic problems. They were also fluent in English and agreed to be available by phone during the study period. Our study population was enrolled as a convenience sample rather than a random sample from the general population. Height and weight were measured and surveys were completed to gather demographic information. Social desirability, or the tendency to avoid criticism and portray one's self in a more favorable manner (46), was measured using two scales. In adolescents, we used the Revised Children's Manifest Anxiety Scale (Lie Scale, or RCMAS-Lie)(39). The 9-item RCMAS-Lie scale was developed specifically for 6 to 19 year olds, has established psychometric properties (39), and has been associated with reporting bias in diet and physical activity in 8 to 10 year old girls enrolled in an intervention (22). In adults, we used the 33-item Marlowe-Crowne Social Desirability Scale that has established psychometric properties (24) and has been linked to reporting biases in diet and physical activity in adults (1;18). Higher scores on both scales indicate higher levels of social desirability. A social desirability bias would be observed if the scales were associated with under-reporting sedentary time and over-reporting physically active time. Informed consent and/or assent was signed by 224 participants (and parents of adolescents), and 213 of these individuals (95%) provided information for the measurements being evaluated. The Institutional Review Board's at Vanderbilt University and the University of Massachusetts approved all study activities.

Previous-day Recall (PDR)

The recall employed was an updated version of our Twenty-four Hour Physical Activity Recall that has been evaluated as a measure of physical activity (8) and used as a reference instrument in an earlier study (27). We use the name Previous-day Recall (PDR) here because, in addition to physical activity, the instrument now gathers more detailed information about sedentary behaviors. Interviewers were certified to complete the recalls using a standard training protocol composed of didactic and experiential training sessions designed to develop interviewing skills, expertise in interacting with the computer interface, and the integration of these two skills. During the study, interviewers led participants chronologically through the previous day (midnight to midnight) using a semi-structured interview based on methods developed and refined for the 7-Day Physical Activity Recall (41). Interviewers gathered information about specific active and sedentary behaviors reported in three segments of the recall day (i.e., morning, afternoon, evening). Individual behaviors lasting at least 5 minutes in a given time-period were recorded/coded and the duration of the activities were entered directly into a database. Each behavior reported was coded as physically active or sedentary using reported body position and activity type (i.e., all exercise and sports pursuits were classified as "active"), and by the location and purpose of the activity. Additional information about the PDR is provided (see Appendix 1, SDC1, Previous Day Recall Protocol). After completing each recall, interviewers assessed the overall reliability of the interview. Interviewers classified recalls as unreliable if the respondent was clearly unable to complete the recall or provide useful information for the

majority of the recall day. Sixteen of the 635 recalls (2.5%) were judged by the interviewers to be unreliable and were excluded from analysis. We defined sedentary behaviors as any behavior that was done while sitting, reclining, or lying down during the waking day, and that did not require substantial energy expenditure (typically < 1.8 metabolic equivalents (METs)) (36). In contrast, physically active behaviors were defined as standing activities, or activities done in any position that resulted in higher MET levels (typically > 1.8 METs). Exercise, sports and active recreation pursuits were classified as active regardless of body position. Each activity in the database was derived from the Compendium of Physical Activities, along with the associated MET values (2). To summarize the recall data, we summed the duration estimates of the individual sedentary and active behaviors that were reported (hrs/d), typically 15 to 30 different activities per recall. For the physically active behaviors, we also calculated time reported in light (< 3 METs), moderate (3–5.9 METs) and vigorous intensity (6+ METs) activity.

Reference Measurement

The activPAL (PAL Technologies, Glasgow, Scotland) is worn on the mid-right thigh, and uses information about thigh position to estimate time spent in different body positions (horizontal = lying or sitting; vertical=standing or stepping). To do so, the instrument records the start and stop time of each individual bout (or event) of lying or sitting, standing, and stepping. Participants wore the device during waking hours, exclusive of bathing and swimming. They were asked to record the time they got out of/into bed and the times they wore the monitor each day. For the activPAL we defined sedentary behavior as time spent sitting or lying during the waking day, and physically active behavior as the sum of time spent standing or stepping. The device also estimates the energy cost of ambulatory activities using a prediction equation that employs stepping cadence and duration as the predictor variables ($\text{MET-hours} = (1.4 \times \text{duration [hours]}) + (4 - 1.4) \times (\text{cadence [steps/minute]}/120) \times \text{duration}$) (37). For descriptive purposes, we also calculated time recorded in moderate-vigorous stepping activities (i.e., 3+ METs). ActivPAL accuracy for measuring body posture in laboratory settings is 95 to 100% (15), and Kozey Keadle (23) reported strong agreement for posture ($R^2=0.94$) between activPAL and direct observation in a free-living study. In an internal validity study, we examined 27 participants over 47 free-living periods of direct observation. Linear mixed models, which included a subject-specific random intercept, revealed a strong linear relationship between the activPAL and direct observation (DO). For sedentary time (min/d) the regression equation was $\text{activPAL} = -1.43 + 1.00 \times \text{DO}$ and for active time (min/d) the regression equation was $\text{activPAL} = 1.41 + 1.00 \times \text{DO}$. The correlation for both measures was $R=0.98$ (unpublished observations).

The ActiGraph (model GT3X) is a triaxial accelerometer that was secured to the right hip using an elastic belt. The monitor was initialized to record vertical acceleration in one-second epochs using the low-frequency extension. Sedentary time was defined as the sum of hours below 100 counts/minute (cpm) and active time was defined as time spent at or above 100 cpm (17;28). Light intensity activity was estimated as time recorded between 100 and 759 cpm, and moderate-vigorous time was estimated using two cut-points. We employed the 760 cpm cut-point that was calibrated to capture a broad range of lifestyle and ambulatory activities with an energy expenditure of 3 METs or greater (26). This cut-point has been cross-validated in free-living studies against indirect calorimetry (26), pattern recognition monitors (50), and time-use diaries (48). We also used the moderate-vigorous cut-point of Freedson (1952 cpm) that was calibrated to capture walking and running behaviors (13) and that has been cross-validated against indirect calorimetry (14;20) an activity diary (43), and other accelerometers (50). Due to the small amount of time recorded in vigorous activity (5725+ cpm), we combined vigorous with moderate activity time for analysis.

Activity Monitor Summary and Wear Time Estimation—To determine monitor wear time for both devices, we used a combination of the wear log information and the automated wear time estimate of Choi (9). The algorithm was set to use any non-zero value of activity counts (ActiGraph) or device movement (activPAL), the time window for consecutive minutes of 0 counts/movement was set at 60 minutes, and the artifact movement detection was set to allow interruptions of 2 minutes or less. Minimum wear time for a “valid” day was 10+ hours. For analysis we calculated estimates of time estimated in sedentary and active behaviors in terms of absolute duration (hrs/d) and as a proportion of total wear time (% wear).

Statistical Methods

Participants eligible for this analysis (N=213) provided 619 valid PDR days, 1,178 valid activPAL days, and 1,277 valid ActiGraph days. We first matched each instrument by date of assessment. Next, because our data collection protocol allowed for shorter valid assessment days for the monitors (minimum 10 hours) compared to the PDR (no minimum), and more than 90% of PDR days had 12+ hours of waking time reported, we also matched each instrument on daily observation time (± 2 hours). Of the 448 PDR-activPAL date matches, 345 days of assessment were within ± 2 hours/day for each method (n=179 participants). From the 1,029 ActiGraph-activPAL date matches, 915 days of assessment were within ± 2 hours/day (n=185 participants). To investigate the possibility that our decision to minimize the impact of extraneous variation in daily observation time between measures of absolute duration by matching on PDR observation and/or monitor wear time could have influenced our results we conducted sensitivity analyses. First, we fitted the measurement error models described below to the 448 days of PDR-activPAL observation not matched on observed/wearing time (n=197) as well as the 1,029 days of unmatched ActiGraph-activPAL data using the absolute duration values (hrs/d). Next, we fitted models for the PDR-activPAL comparisons using % sedentary time estimates (i.e., % observed; % wear) on both the matched and unmatched days.

Measurement Error Modeling

Ideally, we would want to assess the level of agreement between the true, but unobserved, hours of time spent in active and sedentary behaviors on a given day with the corresponding values estimated by the PDR and ActiGraph (AG) instruments. Specifically, we would inquire about the relationships between S_{ij} and S_{ij}^{PDR} , and between S_{ij} and S_{ij}^{AG} , where S_{ij} , S_{ij}^{PDR} , and S_{ij}^{AG} are the hours individual i spent in sedentary behaviors on day j in truth, as estimated by the PDR, and as estimated by the AG. Similarly, we would inquire about the relationships between A_{ij} and A_{ij}^{PDR} , and between A_{ij} and A_{ij}^{AG} , where A_{ij} , A_{ij}^{PDR} , and A_{ij}^{AG} are the hours spent in active behavior in truth, and as estimated by the PDR and AG. For our purposes, as fully explained in the supplementary material (see Appendix 2, Supplemental Digital Content 2, <http://links.lww.com/MSS/A221>, Full Description of Measurement Error Modeling Methods and Assumptions), we chose to treat the activPal measures of sedentary and active behavior, S_{ij}^{AP} and A_{ij}^{AP} respectively as error-free estimates of the truth. Therefore, we model the desired relationships as

$$S_{ij}^T = \beta_0^S + \beta_1^S S_{ij}^{AP} + r_i^S + \varepsilon_{ij}^S \quad \text{equation 1}$$

$$A_{ij}^T = \beta_0^A + \beta_1^A A_{ij}^{AP} + r_i^A + \varepsilon_{ij}^A \quad \text{equation 2}$$

where the general superscript T can be replaced by either PDR or AG, r_i is the person-specific bias, and ϵ_{ij} is the random errors for the test instrument. We further assume that r_i and ϵ_{ij} are independent and normally distributed with mean 0 and variances σ_r^2 and σ_ϵ^2 .

The four parameters describing the quality of the test instrument are β_0 , β_1 , σ_r^2 and σ_ϵ^2 . The intercept, β_0 , and slope β_1 indicate whether the test instrument, on average, correctly estimates the duration of a given behavior. The ideal values of β_0 and β_1 would be 0 and 1, respectively, indicating that the time reported by the test instrument measure is, on average, proportional to the reference instrument. The variance, σ_r^2 , of the person-specific bias (i.e. between-individual variance) measures the magnitude of systematic over- or under-estimation, while the variance, σ_ϵ^2 , of the random error (i.e. within-individual variance) reflects non-systematic or random measurement error. Ideally, both variances should be near 0. To obtain estimates of the desired coefficients, we fitted the models to the individual days of observation using linear mixed models by lmer from the lme4 package in R and calculated standard errors from 1,000 bootstrapped samples. In addition, the mean difference of each participant's average values (i.e., mean of available days), the standard deviation (SD_{dif}) for those differences, and the coefficient of variation for those differences

$CV_{dif} = SD_{dif} / \bar{S}_{\bullet\bullet}^{AP}$, where $\bar{S}_{\bullet\bullet}^{AP}$ is the average value were also compared. Further comparisons of these data were made by the Bland-Altman approach (5) and using Spearman correlations.

In secondary analyses, we evaluated the PDR reports of light and moderate-vigorous intensity activity. To do so, we employed ActiGraph estimates of these metrics as the reference measure using the structure of equation 2, but replaced A_{ij}^{AP} with A_{ij}^{AG} . This approach assumes that the estimate of light and moderate-vigorous activity from the ActiGraph is an unbiased and precise estimate of the truth. Given the uncertainty regarding this assumption in free-living conditions, estimates of the four parameters of interest for the PDR, in this case, could be biased away from their true values. We also report the mean differences and 95% confidence intervals between the PDR and ActiGraph estimates of light and moderate-vigorous activity.

RESULTS

Descriptive characteristics of our study sample are presented in Table 1. The level of agreement between the PDR and activPAL is reported in Table 2, listing both the estimated coefficients for the mixed model and their correlations. Agreement between PDR and activPAL was high in the adults and boys. Among adults the slope of the regression of PDR on activPAL, for both sedentary and active time, were approximately one ($\beta_1 = 0.97$ to 1.13) and the correlation between relevant pairs of measures were high ($\rho = 0.77$ to 0.81). Decomposition of the error variance in the recalls revealed that random errors (σ_ϵ^2) tended to be larger than person-specific biases (σ_r^2). Among boys, slope values were also close to one ($\beta_1 = 0.88$ and 0.96) and the correlations were similar to those for adults ($\rho = 0.75$ and 0.80). Agreement was slightly weaker in girls. Girls had lower slope values ($\beta_1 = 0.64$ and 0.80) and lower correlations ($\rho = 0.52$ and 0.60) in comparison to adults and boys. Girls also had the lowest person-specific bias.

Results comparing the ActiGraph and activPAL are presented in Table 3. For all groups, slope terms were less than one ($\beta_1 = 0.61$ to 0.73). Correlations were high for adults ($\rho = 0.74$ to 0.79) but slightly lower for adolescents ($\rho = 0.57$ to 0.70). Decomposition of the error variance for the ActiGraph measures revealed that device-specific bias (σ_r^2) tended to be larger than random errors (σ_ϵ^2), particularly among males.

T-tests for mean differences and Bland-Altman results are reported in Table 4. Evaluation of PDR versus activPAL revealed no statistically significant mean differences in time spent in active behaviors (all $p \geq 0.16$), but reported sedentary time was greater than activPAL sedentary time in all groups ($p \leq 0.01$). The $CV_{dif}\%$ ranged from 15 to 32% and the limits of agreement were wide. Spearman correlations for the difference scores between measures and the average of both measures were generally positive, and were statistically significant in adults only. Evaluation of mean differences between ActiGraph and activPAL revealed no statistically significant differences in adults, but in adolescents the ActiGraph underestimated sedentary time and overestimated active time (both $p < 0.01$). The $CV_{dif}\%$ ranged from 12 to 35% and the limits of agreement were wide. Spearman correlations between the ActiGraph difference scores tended to be negative, and were statistically significant in women.

Sensitivity analyses of the measurement error models on data not matched on PDR observation or monitor wear time revealed that in unmatched analyses there was a substantial increase in the amount of random error (σ_e^2) and modest reductions of 0.1 to 0.2 units in the slope terms (β_1) and the correlations (ρ) for both the PDR (see Table 1, Supplemental Digital Content 3, <http://links.lww.com/MSS/A218>, Results for Previous-day Recall without matching) and ActiGraph (see Table 2, Supplemental Digital Content 4, <http://links.lww.com/MSS/A219>, Results for ActiGraph without matching) compared to matched analyses presented in Tables 2 and 3, respectively. Evaluation of the PDR-activPAL data for days matched and unmatched on observation time using % sedentary indices, another method to control for differences in observation time, revealed only minimal variation in results for the slope, correlation, and random error terms by matching status (see Table 3, Supplemental Digital Content 5, <http://links.lww.com/MSS/A220>, Results for Previous-day Recall % sedentary time with and without matching).

Correlates of Reporting Errors in Previous-day Recalls (PDR)

The unexplained difference between PDR and activPAL (i.e., residuals) were not significantly correlated with age, gender, BMI, or social desirability in either adults or adolescents. For example, the Spearman correlations between PDR residuals for time reported in active behaviors and BMI (kg/m^2) and social desirability were 0.03 ($p=0.81$) and -0.002 ($p=0.98$) in adults, and 0.01 ($p=0.94$) and -0.05 ($p=0.63$) in adolescents, respectively. Spearman correlations between PDR residuals for time reported in sedentary behaviors and BMI and social desirability were 0.02 ($p=0.85$) and -0.03 ($p=0.79$) in adults, and 0.08 ($p=0.44$) and 0.14 ($p=0.20$) in adolescents, respectively.

Estimates of Light and Moderate-vigorous Physical Activity by Previous-day Recall (PDR)—We also evaluated PDR reported light and moderate-vigorous intensity activity using common ActiGraph cut-points. Comparison of mean differences revealed that PDR reports of light activity tended to be lower than ActiGraph (100–759 cpm), but there were no significant differences in moderate-vigorous activity by PDR and the ActiGraph (760+ cpm) (i.e., the 95% confidence intervals include 0, Figure 1). About 1 to 1.5 hours more moderate-vigorous activity was reported on the PDR than recorded by ActiGraph 1952+ cpm estimates. Evaluation of PDR reported overall, light, and moderate-vigorous activity using measurement error models with the ActiGraph as the reference measure are reported in Table 5. Briefly, for light activity the slope terms were less than 1 ($\beta_1 = 0.34$ to 0.84) and correlations were $\rho = 0.41$ to 0.63 in adults and boys, but lower in girls ($\rho = 0.18$). Using the 760+ cpm moderate-vigorous activity cut-point as reference, among adults and boys the slope terms were approximately 1 ($\beta_1 = 0.88$ to 1.14) and the correlations were $\rho = 0.49$ to 0.63. Both indicators were lower in girls ($\beta_1 = 0.67$ and $\rho = 0.39$). In all models, person-specific bias (σ_p^2) tended to be less than random error (σ_e^2).

DISCUSSION

In this study of free-living adolescents and adults, we found self-report of time spent in physically active and sedentary behaviors by PDR to be strongly correlated with activPAL measures, particularly in adults, and that random reporting errors were larger than person-specific biases. Consistent with our finding of relatively low amounts of person-specific bias, or a person's proclivity to systematically over- or under-report physical activity, we also found no correlation between age, BMI, or social desirability and reporting errors on the PDRs. Notably, validity of the PDR and ActiGraph were comparable to one another as compared to the activPAL for physically active and sedentary behaviors. The PDR also appeared to provide useful estimates of light and moderate-vigorous intensity activity in comparison to commonly used ActiGraph cut-points. Collectively, results from this study indicate that PDR-based estimates of physically active and sedentary time are valid and unbiased measures of time reported in active and sedentary pursuits. PDRs may be a valuable alternative to traditional questionnaire-based measures of these behaviors in future epidemiological studies, particularly those interested in measuring time spent in different body postures (i.e., sitting vs. standing/active), in specific types of behavior, as well as the location and purpose of activity-related behaviors.

In contrast to most physical activity questionnaires that are designed to assess usual activity levels (e.g., past year), which typically have validity coefficients of 0.3 to 0.5 when compared to doubly labeled water or accelerometer-based measures (34;49), we found much higher levels of validity in adults and boys (0.75 to 0.81), and somewhat better results for girls (0.52 to 0.60) for overall time in physically active and sedentary behavior. A number of studies that have examined the validity of various short-term recall approaches are consistent with our results. Hart (16) compared activPAL sitting time to estimates from a physical activity log completed throughout the day and reported a strong correlation ($r=0.87$) and no mean differences between measures. van der Ploeg (48) compared diary-based time-use estimates of non-occupational time on two separate days to the ActiGraph using 100 and 760 cpm cut-points and reported correlations that were similar to our results for sedentary time ($r=0.57, 0.59$), light activity ($r=0.27, 0.39$) and moderate-vigorous activity ($r=0.57, 0.69$). Ridley (40) compared a computer-based previous-day recall to accelerometer counts in youth and, for those 11 years or older, reported correlations of 0.57 for overall physical activity level and 0.41 for moderate-vigorous activity reported on the recall. Calabro (8) compared an earlier version of the PDR to two pattern recognition monitors and reported strong correlations ($r=0.89$ to 0.91) with total energy expenditure (kcal/kg/d) and slightly lower correlations for moderate-vigorous activity ($r=0.57$ to 0.70). Consistent with our findings indicating little correlation between reporting bias and BMI and social desirability for physical activity, Adams (1) reported no evidence of association with reporting bias and these correlates in multiple PDRs in comparison to doubly labeled water in 81 postmenopausal women. The present study extends the finding of an apparent absence of social desirability on the PDR to sedentary behaviors, as well as to men and adolescents. In contrast, Klesges (22) found a positive correlation between physical activity reporting errors and social desirability in African American girls (8–10 years) enrolled in an intervention. Differences between our studies may be due to the physical activity measures employed, cultural factors, additional demands associated with being in an intervention, or the younger age-group in the Klesges report.

In contrast to our findings of small mean differences and the predominance of random reporting errors in self-report, Nusser (35) reported that estimates of total energy expenditure (TEE, kcal/d) derived from a PDR were greater than values estimated by a pattern recognition activity monitor among 171 women. Also using measurement error models, they reported that person-specific biases were three to four times greater than

random error in the recalls. There are several possible explanations for the differences between our studies. First, use of TEE values (kcal/d) as a proxy measure of “physical activity” is susceptible to errors in estimating resting energy expenditure—the major component of TEE. A positive bias in TEE by recall may have been introduced if resting expenditure was estimated as 1 MET because this method overestimates this quantity, and the bias gets larger as BMI level increases (7). This type of bias, unrelated to participant reporting, may have inflated their estimates of person-specific bias. Second, the relatively strict modeling assumption that the reference instrument is unbiased (32) may not have been adequately fulfilled in their study. The reference measure employed has been found to systematically underestimate TEE at higher levels of expenditure (21;44), and it may be that an *underestimate* of TEE by the reference instrument could result in the appearance of larger amounts of person-specific bias in the study (35), regardless of actual reporting accuracy. Additional analyses of these data using different indices of physical activity and considering potential limitations of the available reference instrument would be valuable to further our understanding of the similarities and differences in results for our respective studies.

The present study had several strengths that merit comment. Our study sample was relatively large ($n \sim 180$), it was composed of both adolescents and adults, and we were able to evaluate the validity of the most basic information reported on the PDR (i.e., activity type (active, sedentary) and duration) compared to the activPAL instrument that employed similar definitions of these constructs. Our inclusion of the parallel analysis of the validity of the ActiGraph, an established instrument (28;47), provides a benchmark against which the PDR results can be compared. Results suggest that that the PDR was comparable to the ActiGraph as a measure of physically active and sedentary behaviors with respect to the linear relation, strength of correlation, and low levels of systematic error (person-specific bias) when compared to the activPAL. We also evaluated time reported in light and moderate-vigorous activity in comparison to estimates from the ActiGraph using a cut-point (760 cpm) that was consistent with the scope of the PDR (i.e., assessment of the full range of moderate-vigorous intensity activities). Results provide some evidence for the value of light and moderate-vigorous activity reported on the PDR. The similarity of PDR reported moderate-vigorous activity time with the ActiGraph 760 cpm cut-point values support the accuracy of both instruments for this metric, and this finding is largely consistent with several free-living studies (8;48;50). Our finding that PDR reports of moderate-vigorous activity was greater than estimates derived from the ActiGraph cut-point calibrated only to walking and running (1952 cpm) are also consistent with reports indicating under-recording of moderate-vigorous time by monitors calibrated in this way, in comparison to other accelerometers (50), activity diaries (29;43), and indirect calorimetry (14). An additional strength was our detailed sensitivity analyses to evaluate key aspects of our matching approach to minimizing the known variation in daily observation time between measures. Results suggested that variation in observation time between the instrument being tested and the reference measure was a substantial source of random error that may have exerted a modest negative influence on the slope terms and the correlations between measures and thus supported our decision to employ matching to minimize this source of variability.

There are also limitations to the present study that must be considered. First, our study population was a convenience sample of adolescents and adults that were primarily Caucasian, well-educated and largely working or going to school during their time of study participation. Results may be different in study populations with different demographic characteristics and work and school schedules. Second, our study design limited our ability to investigate possible differences between week- and weekend days using measurement error models. To do so would require replicate measures on both week- and weekend days and we only assessed weekend days once. Future studies should examine this issue more closely since measures of behavior on both types of day may be needed to generate useful

long-term averages (e.g., per week or year)(35). In addition, we did not address the important question regarding the number of replicate PDR measures that are needed to account for seasonal and day of the week effects, as well as true variation in behavior from day-to-day (e.g., (31)). In the context of large-scale epidemiologic investigations we have recently shown that a relatively small number of replicate recalls (e.g., 3 to 4 recalls) obtained using random sampling can substantially reduce the impact of day-to-day variation on behavior-disease associations (32), but more research is needed to enhance our knowledge in this area (e.g., (4)). Another limitation of our study was the lack of an accurate and precise measure of physical activity intensity for use as a reference measurement. While the convergence of the PDR and ActiGraph (760 cpm) estimates for moderate-vigorous activity lends some support for both instruments, uncertainty remains regarding the precision of the ActiGraph measures, and this source of error could underestimate the apparent validity of the PDR. Indeed, we observed that the correlations with overall active time were 15 to 25% lower, and the variance estimates for person-specific bias and random error were larger, when the ActiGraph was used to evaluate the PDR (see Table 5), compared to results using the activPAL for reference (see Table 2). Thus, use of the ActiGraph in this context may modestly underestimate the validity of PDR reported light and moderate-vigorous activity. Clearly, future studies are needed to extend our understanding of the validity of reports for different activity intensities using better reference measures, as well as the validity of the contextual information reported on the PDR (i.e., location, purpose of activity).

Conclusion

This report provides proof of principle that the PDR may be a valid method for measurement of physically active and sedentary behaviors in epidemiologic studies that seek to rank-order individuals by level of a given behavioral exposure. PDRs may be particularly useful for studies that desire to assess the full range of human behavior, body position, and also gather details about where and why these behaviors occur. Future studies are needed to replicate these findings among larger, more ethnically diverse study populations, and to evaluate the ability of the interviewer-based PDRs to be translated to self-administered PDRs suitable for large scale studies (e.g. internet-based instruments, mobile devices).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank Cara Hanby, Mary Kay Fadden, Stacey Peterson, and Sara Hollis for their integral work in helping develop and refine the PDR method and the initial infrastructure for the present study.

Financial Support

This research was supported by funding from the National Institutes of Health (R01NR011477) to Drs. Fowke and Freedson. The Intramural Research Program of the National Institutes of Health supported Dr. Matthews work on this project.

Reference List

1. Adams SA, Matthews CE, Moore CG, Cunningham JE, Fulton J, Hebert JR. The effect of social desirability and social approval on self-reports of physical activity. *American Journal of Epidemiology*. 2005; 161(4):389–98. [PubMed: 15692083]
2. Ainsworth B, Haskell W, Whitt M, Irwin M, AS, Strath S, O'Brien W, Bassett D, Schmitz K, Emplainscourt P, et al. Compendium of Physical Activities: An Update of Activity Codes and MET

- Intensities. *Medicine & Science in Sports & Exercise*. 2000; 32(9):S498–S516. [PubMed: 10993420]
3. Atkin AJ, Gorely T, Clemes SA, Yates T, Edwardson C, Brage S, Salmon J, Marshall SJ, Biddle SJ. Methods of Measurement in epidemiology: Sedentary Behaviour. *Int J Epidemiol*. 2012 Oct 1; 41(5):1460–71. [PubMed: 23045206]
 4. Baranowski T, Masse LC, Ragan BWG. How Many Days Was That? We're Still Not Sure, But We're Asking the Question Better! *Medicine & Science in Sports & Exercise*. 2008; 40(Suppl 7):S544–9. [PubMed: 18562972]
 5. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 1(8476):307–10. [PubMed: 2868172]
 6. Bowles HR. Measurement of Active and Sedentary Behaviors: Closing the Gaps in Self-Report Methods. *Journal of Physical Activity and Health*. 2012; 9(Suppl 1):S1–S4. [PubMed: 22287442]
 7. Byrne NM, Hills AP, Hunter GR, Weinsier RL, Schutz Y. Metabolic equivalent: one size does not fit all. *Journal of Applied Physiology*. 2005 Sep 1; 99(3):1112–9. [PubMed: 15831804]
 8. Calabro MA, Welk GJ, Carriquiry AL, Nusser SM, Beyler NK, Matthews CE. Validation of a Computerized 24-Hour Physical Activity Recall (24PAR) Instrument With Pattern-Recognition Activity Monitors. *Journal of Physical Activity and Health*. 2009; 6:211–20. [PubMed: 19420399]
 9. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of Accelerometer Wear and Nonwear Time Classification Algorithm. *Medicine & Science in Sports & Exercise*. 2011; 43:357–64. [PubMed: 20581716]
 10. Ferrari P, Friedenreich C, Matthews CE. The Role of Measurement Error in Estimating Levels of Physical Activity. *American Journal of Epidemiology*. 2007 Oct 1; 166(7):832–40. [PubMed: 17670910]
 11. Freedman LS, Schatzkin A, Midthune D, Kipnis V. Dealing With Dietary Measurement Error in Nutritional Cohort Studies. *J Natl Cancer Inst*. 2011 Jun 8; 103:1086–92. [PubMed: 21653922]
 12. Freedson PS, Bowles HR, Troiano R, Haskell WL. Assessment of Physical Activity Using Wearable Monitors: Recommendations for Monitor Calibration and Use in the Field. *Medicine & Science in Sports & Exercise*. 2012; 44(1S):S1–S4. [PubMed: 22157769]
 13. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine & Science in Sports & Exercise*. 1998; 30:777–81. [PubMed: 9588623]
 14. Freedson PS, Lyden K, Kozey-Keadle S, Staudenmayer J. Evaluation of artificial neural network algorithms for predicting METs and activity type from accelerometer data: validation on an independent sample. *Journal of Applied Physiology*. 2011 Dec 1; 111(6):1804–12. [PubMed: 21885802]
 15. Grant PM, Ryan CG, Tigbe WW, Granat MH. The validation of a novel activity monitor in the measurement of posture and motion during everyday activities. *British Journal of Sports Medicine*. 2006 Dec 1; 40(12):992–7. [PubMed: 16980531]
 16. Hart TL, Ainsworth BE, Tudor-Locke C. Objective and Subjective Measures of Sedentary Behavior and Physical Activity. *Medicine & Science in Sports & Exercise*. 2011; 43(3):449–56. [PubMed: 20631642]
 17. Healy GN, Clark B, 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):216–27. [PubMed: 21767730]
 18. Hebert JR, Ebbeling CB, Matthews CE, Ma Y, Clemow L, Hurley TG, Druker SK, Rosal MC, Ockene JK. Social Desirability and Approval-Related Biases in Middle-Aged Women's Estimates of Energy Intake: Comparing Structured Dietary Questionnaires to Total Energy Expenditure from Doubly Labeled Water. *Annals of Epidemiology*. 2002; 12:577–86. [PubMed: 12495831]
 19. Helmerhorst HJ, Brage S, Warren J, Besson H, Ekelund U. A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. *International Journal of Behavioral Nutrition and Physical Activity*. 2012; 9(103):1–55. [PubMed: 22233712]
 20. Hendelman D, Miller K, Baggett C, Debold E, Freedson P. Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Medicine & Science in Sports & Exercise*. 2000; 32(9 Suppl):442–9.

21. Johannsen DL, Calabro MA, Stewart J, Franke W, Rood JC, Welk GJ. Accuracy of Armband Monitors for Measuring Daily Energy Expenditure in Healthy Adults. [Miscellaneous Article]. *Medicine & Science in Sports & Exercise*. 2010 Nov; 42(11):2134–40. [PubMed: 20386334]
22. Klesges LM, Baranowski T, Beech B, Cullen K, Murray DM, Rochon J, Pratt C, Klesges LM, Baranowski T, Beech B, et al. Social desirability bias in self-reported dietary, physical activity and weight concerns measures in 8- to 10-year-old African-American girls: results from the Girls Health Enrichment Multisite Studies (GEMS). *Preventive Medicine*. 2004 May; 38(Suppl):S78–S87. [PubMed: 15072862]
23. Kozey-Keadle S, Libertine A, Lyden K, Staudenmayer J, Freedson P. Validation of Wearable Monitors for Assessing Sedentary Behavior. *Medicine & Science in Sports & Exercise*. 2011; 43:1561–7. [PubMed: 21233777]
24. Marlowe D, Crowne DP. Social desirability and response to perceived situational demands. *Journal of Consulting Psychology*. 1961; 25(2):109–15. [PubMed: 13766997]
25. Matthews, CE. Techniques for Physical Activity Assessment: Self-Report Instruments. In: Welk, G.; Dale, D., editors. *Physical Activity Assessments for Health-Related Research*. Champaign, IL: Human Kinetics; 2002. p. 107-23.
26. Matthews CE. Calibration of Accelerometer Output for Adults. *Med Sci Sports Exerc*. 2005; 37(11):S512–S522. [PubMed: 16294114]
27. Matthews CE, Ainsworth BE, Hanby C, Pate RR, Addy C, Freedson PS, Jones DA, Macera CA. Development and testing of a short physical activity recall questionnaire. *Medicine & Science in Sports & Exercise*. 2005 Jun; 37(6):986–94. [PubMed: 15947724]
28. Matthews CE, Chen KY, Freedson PS, Buchowski MS, Beech BM, Pate RR, Troiano RP. Amount of time spent in sedentary behaviors - United States 2003–2004. *American Journal of Epidemiology*. 2008; 167:875–81. [PubMed: 18303006]
29. Matthews CE, Freedson PS. Field trial of a three-dimensional activity monitor: comparison with self-report. *Medicine & Science in Sports & Exercise*. 1995; 27(7):1071–8. [PubMed: 7564975]
30. Matthews CE, George SM, Moore SC, Bowles HR, Blair A, Park Y, Troiano RP, Hollenbeck AR, Schatzkin A. Amount of time spent in sedentary behaviors and cause-specific mortality in US adults. *American Journal of Clinical Nutrition*. 2012; 95:437–45. [PubMed: 22218159]
31. Matthews CE, Hebert JR, Freedson PS, EJ, Stanek I, Merriam PA, Ebbeling CB, Ockene IS. Sources of Variance in Daily Physical Activity Levels in the Seasonal Variation of Blood Cholesterol Study. *American Journal of Epidemiology*. 2001; 153(10):987–95. [PubMed: 11384955]
32. Matthews CE, Moore SC, George SM, Sampson J, Bowles HR. Improving Self-reports of Active and Sedentary Behaviors in Large Epidemiologic Studies. *Exercise and Sports Science Reviews*. 2012; 40(3):118–26.
33. Moore SC, Gierach GL, Schatzkin A, Matthews CE. Physical activity, sedentary behaviours, and the prevention of endometrial cancer. *Br J Cancer*. 2010 Sep 28; 103(7):933–8. [PubMed: 20877336]
34. Neilson HK, Robson PJ, Friedenreich CM, Csizmadia I. Estimating activity energy expenditure: how valid are physical activity questionnaires? *American Journal of Clinical Nutrition*. 2008 Feb 1; 87(2):279–91. [PubMed: 18258615]
35. Nusser SM, Beyler NK, Welk GJ, Carriquiry AL, Fuller WA, King BMN. Modeling Errors in Physical Activity Recall Data. *Journal of Physical Activity and Health*. 2012; 9(Suppl 1):S56–S67. [PubMed: 22287449]
36. Owen N, Sparling PB, Healy GvN, Dunstan DW, Matthews CE. Sedentary Behavior: Emerging Evidence for a New Health Risk. *Mayo Clinic Proceedings*. 2010 Dec 1; 85(12):1138–41. [PubMed: 21123641]
37. PAL Technologies. Appendix A - Technical Description. 2010. *activPAL Operating Guide*; p. 15-17.
38. Physical Activity Guidelines Advisory Committee. *Physical Activity Guidelines Advisory Committee Report*. Washington, DC: U.S. Department of Health and Human Services; 2008. p. A1-A10.

39. Reynolds CR, Paget KR. National normative and reliability data for the revised children's manifest anxiety scale. *School Psych Rev.* 1983; 12:324–36.
40. Ridley K, Olds T, Hill A. The Multimedia activity recall for children and adolescents (MARCA): development and evaluation. *International Journal of Behavioral Nutrition and Physical Activity.* 2006; 3(1):10. [PubMed: 16725055]
41. Sallis, JF. A collection of physical activity questionnaires for health-related research: Seven-day physical activity recall. In: Kriska, AM.; Casperson, CJ., editors. *A Collection of Physical Activity Questionnaires.* Medicine & Science in Sports & Exercise. 1997. p. S89-S103.
42. Schatzkin A, Subar AF, Moore S, Park Y, Potischman N, Thompson FE, Leitzmann M, Hollenbeck A, Morrissey KG, Kipnis V. Observational Epidemiologic Studies of Nutrition and Cancer: The Next Generation (with Better Observation). *Cancer Epidemiology Biomarkers & Prevention.* 2009 Apr; 18(4):1026–32.
43. Sirard JR, Melanson EL, Li L, Freedson PS. Field evaluation of the Computer Science and Applications, Inc. physical activity monitor. *Medicine & Science in Sports & Exercise.* 2000; 32(3):695–700. [PubMed: 10731015]
44. St-Onge M, Mignault D, Allison DB, Rabasa-Lhoret R+ Evaluation of a portable device to measure daily energy expenditure in free-living adults. *The American Journal of Clinical Nutrition.* 2007 Mar 1; 85(3):742–9. [PubMed: 17344495]
45. Thorp AA, Owen N, Neuhaus M, Dunstan DW. Sedentary Behaviors and Subsequent Health Outcomes in Adults: A Systematic Review of Longitudinal Studies, 1996–2011. *American Journal of Preventive Medicine.* 2011 Aug; 41(2):207–15. [PubMed: 21767729]
46. Tourangeau, R.; Rips, LJ.; Rasinski, K. Editing of Responses: Reporting About Sensitive Topics. In: Tourangeau, R.; Rips, LJ.; Rasinski, K., editors. *The Psychology of Survey Response.* 1. Cambridge: Cambridge University Press; 2000. p. 257-8.
47. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical Activity in the United States Measured by Accelerometer. *Medicine & Science in Sports & Exercise.* 2008; 40(1): 181–8. [PubMed: 18091006]
48. van der Ploeg HP, Merom D, Chau JY, Bittman M, Trost SG, Bauman AE. Advances in Population Surveillance for Physical Activity and Sedentary Behavior: Reliability and Validity of Time Use Surveys. *American Journal of Epidemiology.* 2010 Nov 15; 172(10):1199–206. [PubMed: 20855469]
49. van Poppel MNM, Chinapaw MJM, Mokkink LB, van Mechelen W, Terwee CB. Physical Activity Questionnaires for Adults: A Systematic Review of Measurement Properties. *Sports Medicine.* 2010; 40(7):565–600. [PubMed: 20545381]
50. Welk GJ, McClain JJ, Eisenmann JC, Wickel EE. Field Validation of the MTI Actigraph and BodyMedia Armband Monitor Using the IDEEA Monitor. *Obesity.* 2007 Apr 1; 15(4):918–28. [PubMed: 17426327]

Figure 1.

Difference between Previous-Day Recall (PDR) and ActiGraph measures (hrs/d) of light and moderate-vigorous intensity physical activity (n=183)¹, by age-group and gender.

Values are Mean and 95% confidence intervals

¹ Data are derived from 2.2 (SD=0.8) days of assessment per participant and 14.3 (SD=1.) hours of ActiGraph wear time and 14.7 (SD=1.6) hours of PDR waking time.

Table 1

Characteristics of Study Participants

	Adolescents	Adults
Age (years)	14.3 (1.7)	41.3 (14.8)
Body mass index (kg/m²)	22.0 (5.8)	26.9 (5.4)
Body mass index 30+ kg/m² (%)	6.7%	22.3%
Female (%)	49.1%	53.7%
Education (%)		
6–8th grade	53.1%	
9–12th grade	44.8%	1.1%
High School Graduate	2.1%	9.1%
Some college		26.1%
Bachelor's degree		30.7%
Graduate degree		33.0%
Race/ethnicity		
White	61.3%	74.7%
Black	16.0%	14.7%
Other	13.2%	3.2%
Unknown	9.4%	7.4%
In school during study (%)	59.4%	
Working during study (%)	20.8%	83.0%
Social Desirability	2.7 (2.4)	17.5 (2.8)
Previous-day Recalls (PDR)		
Waking time (hrs/d)	14.2 (1.47)	15.1 (1.50)
Sedentary (hrs/d)	9.9 (2.23)	10.1 (2.91)
Active (hrs/d)	4.3 (1.81)	5.1 (2.58)
Light activity (hrs/d, < 3 METS) ¹	2.3 (1.23)	3.4 (1.79)
Moderate activity (hrs/d, 3–5.9 METS) ¹	1.5 (1.13)	1.5 (1.63)
Vigorous activity (hrs/d, 6+ METS) ¹	0.6 (0.83)	0.2 (0.35)
ActivPAL		
Wear time (hrs/d)	13.3 (1.13)	14.4 (1.17)
Sedentary (hrs/d, Sit/lie)	9.0 (1.44)	9.0 (1.90)
Active (hrs/d, Upright)	4.3 (1.34)	5.5 (1.87)
Standing still (hrs/d)	2.8 (1.05)	3.8 (1.47)
Stepping (hrs/d)	1.6 (0.53)	1.6 (0.59)
Stepping, Mod-Vig (hrs/d, 3+ METS) ²	0.7 (0.30)	0.7 (0.36)
ActiGraph		
Wear time (hrs/d)	13.6 (1.11)	14.8 (1.17)
Sedentary (hrs/d, < 100 cpm)	7.7 (1.32)	9.1 (1.71)
Active (hrs/d, 100+ cpm)	5.9 (1.29)	5.7 (1.64)
Light activity (hrs/d, 100–759 cpm) ³	3.7 (0.70)	3.7 (1.00)

	Adolescents	Adults
Moderate activity (hrs/d, 760–5724 cpm) ³	2.0 (0.71)	1.9 (0.88)
Vigorous activity (hrs/d, 5725+ cpm) ³	0.04 (0.11)	0.04 (0.08)
Freedson Mod-Vig (hrs/d, 1952+ cpm) ³	0.7 (0.44)	0.6 (0.40)

Values are mean (SD) and percentages (%); cpm = activity counts per minute; hrs/d = hours per day; kg/m² = kilograms per meter squared; METS = metabolic equivalents; MET-hrs/d = metabolic equivalent hours per day; Mod-Vig = moderate-vigorous

¹ Active hours classified by intensity: Light (< 3 METs); Moderate (3–5.9 METs), Vigorous intensity (6+ METs)

² Stepping, Mod-Vig = sum of hours stepping at or above moderate-vigorous intensity (3+ METs)

³ Hours spent in light (100–759 cpm) and vigorous activity (5725+ cpm), and for two moderate intensity cut-points (1952–5724 cpm (12)) and (760–5724 cpm (23)).

Results Comparing Previous-day Recall Measures of Sedentary and Active time (hrs/d) to the activPAL (n=179)¹, by Age-group and Gender

Table 2

Group	Type of behavior	Model parameters/results					Correlation ρ
		Intercept β_0	Slope β_1	Total Variance σ_{AP}^2	Person-specific σ_r^2	Random error σ_ε^2	
ADULTS							
Men (n=40)	Sedentary	-0.55 (1.06)	1.13 (0.12)	5.46 (0.91)	0.58 (0.49)	3.43 (0.57)	0.81 (0.05)
	Active	0.16 (0.77)	0.97 (0.13)	4.74 (1.11)	0.78 (0.65)	2.79 (0.51)	0.77 (0.08)
Women (n=48)	Sedentary	0.27 (0.67)	1.05 (0.07)	5.40 (0.68)	1.24 (0.52)	1.89 (0.31)	0.81 (0.04)
	Active	-0.67 (0.41)	1.09 (0.08)	4.43 (0.62)	0.91 (0.42)	2.02 (0.38)	0.80 (0.05)
ADOLESCENTS							
Boys (n=43)	Sedentary	1.18 (0.83)	0.96 (0.09)	6.47 (1.23)	0.70 (0.49)	2.67 (0.55)	0.80 (0.05)
	Active	0.27 (0.46)	0.88 (0.09)	4.03 (1.06)	0.44 (0.40)	2.10 (0.47)	0.75 (0.07)
Girls (n=48)	Sedentary	2.52 (0.84)	0.80 (0.09)	3.34 (0.47)	0.27 (0.47)	3.51 (0.55)	0.60 (0.06)
	Active	1.47 (0.48)	0.64 (0.09)	2.67 (0.39)	0.21 (0.41)	2.82 (0.44)	0.52 (0.07)

Standard errors are in parentheses and were estimated from 1,000 bootstrapped samples

¹Data are derived from 1.9 (SD=0.8) days per participant and 14.7 (SD=1.5) hours of PDR waking time, and 14.1 (SD=1.5) hours of activPAL wear time.

Table 3
Results Comparing ActiGraph Measures of Sedentary and Active Time (hrs/d) to the activPAL (n=185)¹, by Age-group and Gender

Group	Type of behavior	Model parameters/results					
		Intercept β_0	Slope β_1	Total Variance σ_{AP}^2	Person-specific σ_r^2	Random error σ_ε^2	Correlation ρ
ADULTS							
Men (n=40)	Sedentary	2.27 (0.61)	0.73 (0.05)	6.20 (1.09)	1.41 (0.46)	0.78 (0.10)	0.75 (0.07)
	Active	2.30 (0.30)	0.65 (0.07)	5.24 (1.08)	1.29 (0.37)	0.67 (0.10)	0.74 (0.08)
Women (n=44)	Sedentary	3.65 (0.44)	0.63 (0.04)	5.40 (0.67)	0.83 (0.19)	0.78 (0.10)	0.77 (0.04)
	Active	2.23 (0.27)	0.61 (0.05)	5.47 (0.67)	0.61 (0.15)	0.65 (0.10)	0.79 (0.04)
ADOLESCENTS							
Boys (n=51)	Sedentary	1.22 (0.47)	0.69 (0.05)	4.85 (0.72)	1.27 (0.32)	0.87 (0.12)	0.68 (0.06)
	Active	3.09 (0.31)	0.72 (0.07)	3.53 (0.83)	1.35 (0.32)	0.82 (0.11)	0.57 (0.07)
Girls (n=50)	Sedentary	1.99 (0.44)	0.67 (0.05)	3.18 (0.31)	0.76 (0.18)	0.70 (0.08)	0.70 (0.04)
	Active	2.72 (0.21)	0.68 (0.04)	2.24 (0.21)	0.64 (0.12)	0.55 (0.07)	0.67 (0.05)

Standard errors are in parentheses and were estimated from 1,000 bootstrapped samples

¹Data are derived from 4.9 (SD=1.7) days of assessment per participant and 14.2 (SD=1.2) hours of ActiGraph and 13.9 (SD=1.3) hours of activPAL wear time.

Table 4
Evaluation of mean differences between measures (hrs/d), Bland-Altman analysis, by age-group and gender

<i>Previous-day Recall (PDR) vs. activPAL</i>						Bland-Altman Analysis			
		activPAL	PDR	Mean difference	SD difference	CV _{diff} %	P (t- test)	Limits of Agreement (hrs/d)	Spearman Correlation ^{**}
Adults									
Men (n=40)	Sedentary	9.16 (2.09)	9.88 (3.21)	0.72	1.70	19%	0.01	-2.61	4.05
	Active	5.44 (2.13)	5.39 (2.84)	-0.05	1.58	29%	0.83	-3.15	3.05
Female (n=48)	Sedentary	9.19 (1.91)	9.94 (2.51)	0.75	1.51	16%	< 0.01	-2.21	3.71
	Active	5.27 (1.93)	5.08 (2.38)	-0.20	1.32	25%	0.31	-2.79	2.39
Adolescents									
Boys (n=43)	Sedentary	9.36 (2.17)	10.19 (2.56)	0.82	1.44	15%	< 0.01	-2.00	3.64
	Active	4.44 (1.85)	4.20 (2.04)	-0.24	1.27	29%	0.22	-2.73	2.25
Girls (n=48)	Sedentary	9.01 (1.75)	9.81 (1.93)	0.81	1.60	18%	< 0.01	-2.33	3.95
	Active	4.56 (1.49)	4.26 (1.62)	-0.30	1.46	32%	0.16	-3.17	2.57
ActiGraph vs. activPAL									
		activPAL	Actigraph	Mean difference	SD difference	CV%	P (t- test)	Limits of Agreement (hrs/d)	Spearman Correlation ^{**}
Adults									
Men (n=40)	Sedentary	9.19 (2.14)	8.98 (1.85)	-0.22	1.47	16%	0.36	-3.10	2.66
	Active	5.48 (1.93)	5.87 (1.79)	0.39	1.35	25%	0.08	-2.26	3.04
Female (n=44)	Sedentary	8.83 (1.76)	9.16 (1.61)	0.32	1.12	13%	0.06	-1.88	2.52
	Active	5.60 (1.94)	5.64 (1.48)	0.04	1.15	21%	0.83	-2.21	2.29
Adolescents									
Boys (n=51)	Sedentary	9.14 (1.62)	7.48 (1.42)	-1.66	1.44	16%	< 0.01	-4.48	1.16
	Active	4.11 (1.51)	6.04 (1.41)	1.93	1.45	35%	< 0.01	-0.91	4.77
Girls (n=50)	Sedentary	8.79 (1.14)	7.92 (1.21)	-0.87	1.06	12%	< 0.01	-2.95	1.21
	Active	4.54 (1.07)	5.81 (1.10)	1.27	0.98	22%	< 0.01	-0.65	3.19

SD difference= standard deviation of difference score; CV_{diff}%=SD difference/activPAL mean; Spearman=correlation of difference scores and average of both measures;

* p < 0.05 Spearman Correlation

** correlation between residual difference between measures and the mean of both measures

Table 5

Results comparing Previous-day Recall measures of overall active time and activity intensity (hrs/d) to the ActiGraph (n=183)¹, by age-group and gender

Group	Type of behavior	Model Results					
		Intercept β_0	Slope β_1	Total Variance σ_{AG}^2	Person-specific σ^2_r	Random error σ^2_e	Correlation ρ
ADULTS							
Men (n=39)	Active	0.02 (0.93)	0.90 (0.15)	4.18 (0.84)	1.90 (0.79)	3.28 (1.05)	0.64 (0.10)
	Light ¹	0.80 (0.64)	0.62 (0.19)	1.49 (0.30)	0.76 (0.56)	2.73 (0.83)	0.41 (0.08)
	Mod-vig-760 ²	0.29 (0.53)	0.88 (0.24)	1.44 (0.43)	1.22 (0.49)	2.29 (0.76)	0.51 (0.15)
	Mod-vig-1952 ³	1.76 (0.37)	0.63 (0.40)	0.34 (0.07)	2.11 (0.79)	2.45 (0.76)	0.22 (0.13)
Women (n=46)	Active	-1.25 (0.62)	1.16 (0.12)	3.09 (0.69)	1.95 (0.61)	2.97 (0.61)	0.68 (0.06)
	Light ²	0.41 (0.66)	0.84 (0.19)	1.50 (0.21)	2.27 (0.82)	2.24 (0.46)	0.45 (0.09)
	Mod-vig-760 ³	-0.41 (0.45)	1.14 (0.27)	0.82 (0.17)	1.14 (0.70)	2.29 (0.61)	0.49 (0.12)
	Mod-vig-1952 ⁴	0.99 (0.29)	1.14 (0.37)	0.27 (0.07)	1.60 (1.02)	2.58 (0.71)	0.26 (0.09)
ADOLESCENTS							
Boys (n=46)	Active	0.54 (1.09)	0.66 (0.19)	2.97 (0.47)	2.24 (0.83)	2.69 (0.72)	0.40 (0.12)
	Light ¹	0.17 (1.24)	0.56 (0.37)	1.23 (0.25)	0.50 (0.51)	2.93 (0.77)	0.29 (0.18)
	Mod-vig-760 ²	-0.11 (0.36)	1.04 (0.14)	1.27 (0.31)	0.51 (0.49)	1.92 (0.68)	0.62 (0.10)
	Mod-vig-1952 ³	0.81 (0.24)	1.72 (0.25)	0.49 (0.09)	0.36 (0.35)	2.01 (0.61)	0.63 (0.08)
Girls (n=52)	Active	1.15 (0.63)	0.55 (0.11)	2.78 (0.35)	1.25 (0.69)	2.55 (0.45)	0.39 (0.09)
	Light ¹	1.36 (0.70)	0.34 (0.18)	1.07 (0.13)	0.12 (0.27)	3.47 (0.61)	0.18 (0.10)
	Mod-vig-760 ²	0.42 (0.22)	0.67 (0.12)	0.87 (0.14)	0.00 (0.09)	2.20 (0.58)	0.39 (0.07)
	Mod-vig-1952 ³	1.26 (0.22)	0.81 (0.23)	0.30 (0.07)	0.00 (0.10)	2.40 (0.65)	0.27 (0.10)

Standard errors are in parentheses and were estimated from 1,000 bootstrapped samples

¹ Data are derived from 2.2 (SD=0.8) days of assessment per participant and 14.3 (SD=1.1) hours of ActiGraph and 14.7 (SD=1.6) hours of PDR waking time.

² Light intensity activity by ActiGraph (100–759 cpm);

³ Moderate-vigorous intensity activity by ActiGraph (760+ cpm);

⁴ Moderate-vigorous intensity activity by ActiGraph (1952+ cpm)