Estimating Minutes of Physical Activity From the Previous Day Physical Activity Recall: Validation of a Prediction Equation

Jared M. Tucker, Greg Welk, Sarah M. Nusser, Nicholas K. Beyler, and David Dzewaltowski

**Background:** This study was designed to develop a prediction algorithm that would allow the Previous Day Physical Activity Recall (PDPAR) to be equated with temporally matched data from an accelerometer. **Methods:** Participants (n = 121) from a large, school-based intervention wore a validated accelerometer and completed the PDPAR for 3 consecutive days. Physical activity estimates were obtained from PDPAR by totaling 30-minute bouts of activity coded as ≥4 METS. A regression equation was developed in a calibration sample (n = 91) to predict accelerometer minutes of moderate to vigorous physical activity (MVPA) from PDPAR bouts. The regression equation was then applied to a separate, holdout sample (n = 30) to evaluate the utility of the prediction algorithm. **Results:** Gender and PDPAR bouts accounted for 36.6% of the variance in accelerometer MVPA. The regression model showed that on average boys obtain 9.0 min of MVPA for each reported PDPAR bout, while girls obtain 4.8 min of MVPA per bout. When applied to the holdout sample, predicted minutes of MVPA from the models showed good agreement with accelerometer minutes (r = .81). **Conclusions:** The prediction equation provides a valid and useful metric to aid in the interpretation of PDPAR results.

**Keywords:** accelerometry, PDPAR, self-report, youth

Accurate measures of physical activity (PA) in children are important for the assessment of current activity levels in youth and to determine how these levels are changing over time. Valid PA measurement tools are also needed to advance research on the health consequences of PA, and PA correlates, as well as to effectively evaluate interventions designed to promote PA. However, obtaining accurate measures of PA in youth can be difficult, especially in large samples. Objective assessments such as indirect calorimetry, heart rate monitoring, or accelerometry are generally favored to improve precision but these tools are often impractical for large populations due to logistics and cost.

Despite limitations, self-report instruments remain the most widely used methodology for PA assessment. There are a wide array of tools available, but single day recall instruments have been most common since there is a reduced cognitive demand of recalling a single day. One of the most widely used recall tools for youth is the Previous Day Physical Activity Recall (PDPAR). This tool is designed to capture detailed information about activity the previous day, but a related instrument called the 3DPAR is available for assessing after school activity on 3 consecutive days. The PDPAR and 3DPAR have been used in a variety of research applications including evaluation of correlates, interventions, and for descriptive epidemiology research. The PDPAR uses a time grid that requires youth to report the predominant type and intensity of the main activity that they performed for each 30 minute period in the day. It was developed primarily to quantify after-school PA in youth, and can be scored to provide information on the intensity, frequency, duration, and mode of physical activities.

A number of studies have evaluated the validity of the PDPAR instrument and results have generally shown it to be psychometrically sound. The original validation study of the PDPAR reported correlations of $r = .77$ when PDPAR energy expenditure estimates were compared with accelerometry, and $r = .63$ (when PDPAR blocks of PA were compared with estimates based on heart rate monitoring). Other studies comparing PA measures from accelerometers with the PDPAR, 3DPAR, and similar time-based recalls, have yielded correlations ranging from 0.19 to 0.57. The variation in these associations appears to be dependent on the PDPAR outcome measures being compared (eg, METs vs. blocks), the amount of time being recalled, and the characteristics of the population sampled. For example, Pate et al...
suggested that the relatively high correlations seen in the initial PDPAR validation might have been due, in part, to a shorter recall period (one afternoon) and a sample that included relatively older (high school) participants. While the PDPAR has been well validated, and has demonstrated utility for various research applications, it also has some limitations. One such drawback is that physical activities are reported in 30-minute time blocks, which may not reflect the sporadic activity patterns of youth. Children are asked to report the predominant activity they performed for each time block, but it may be difficult for children to characterize their activity into one of the specific categories. Another limitation is that the outcome measure of “blocks” of activity is difficult to interpret. One study multiplied the number of blocks by 30 minutes to estimate minutes of PA, but direct conversions to estimates of time are not recommended since children may not be active for a full 30-minute block. Since public health goals regarding appropriate PA levels among youth are based on minutes of PA, it would be advantageous to be able to convert PDPAR results to this unit of measure. There are numerous examples in the exercise science literature of equations that have been developed to allow estimation of one parameter from related sources of data. For example, skinfold measures are routinely used to estimate body fatness, since body fat levels can be readily interpreted from skinfold thicknesses. This type of calibration would have the same utility for PA assessments since it would allow self-report data (typically assumed to be biased) to be converted into a more meaningful measure (ie, minutes of MVPA).

The purpose of the current study was to develop and evaluate prediction algorithms that would allow the amount of MVPA to be estimated from reported PDPAR blocks. The study employed a large sample of participants who had matched data from the PDPAR instrument and an objective PA monitor over a 3-day period (3 single day PDPAR administrations). The algorithms were developed on 1 sample and then cross-validated on an independent sample to evaluate the utility of the conversion algorithms.

Methods

Participants

The study was conducted as part of a large school-based intervention study (Healthy Youth Places) that employed the PDPAR as the primary outcome measure. A subsample of participants from 9 of the 16 schools participated in a supplemental activity monitoring protocol that involved wearing an accelerometer across the same days assessed by the PDPAR. The data are from year 2 of the project and include 7th grade students (n = 240) from both intervention and control schools. Since the primary interest was in the relationship between the 2 temporally matched activity measures, the group designation was not important for the current study. To facilitate evaluation, the sample was randomly divided into a calibration sample (75% of participants) and a separate validation sample (25% of participants). The sample was divided in this manner to provide a larger sample size for the calibration analyses, and therefore, more robust prediction equations. Parental consent and student assent documents were obtained for all participants. The study was approved by the Kansas State University institutional review board and the school district’s research committee.

Instruments

Instruments used in the current investigation included the PDPAR, a PA recall questionnaire, and the Biotrainer accelerometer (IM Systems, Baltimore, MD). Details about the use of the PDPAR and Biotrainer were described in a previous study, using baseline data from this project. Information specific to the current study are described below.

PDPAR

Participants completed the PDPAR on 3 consecutive weekdays, recalling after-school physical activities from the previous day on each occasion. The PDPAR utilizes a time-based recall strategy by dividing the day’s after-school activities into 30-minute time blocks. The dominant activity for each time block is chosen from a list of activities, after which the corresponding intensity level for that activity is chosen from 4 options: very light, light, moderate, or vigorous. Metabolic equivalents (METs) are assigned to each block based on the activity and intensity reported. In the current study, total blocks of MVPA per day (≥4 METs) from the PDPAR was the outcome variable of interest.

Biotrainer Accelerometer

Objective PA assessments were obtained using the Biotrainer accelerometer on the same 3 days as the PDPAR assessments. The Biotrainer is a biaxial accelerometer that is similar to other accelerometry-based monitors (eg, the Actigraph). The Biotrainer uses a high-speed sampling method to accumulate counts over a user-specified interval. Research validating the Biotrainer against indirect calorimetry showed correlations of 0.88 and 0.59 for laboratory and field-based conditions, respectively. The monitor was highly correlated with the Actigraph and the Tritrac for both locomotor (r = .86–.95) and lifestyle activities (r = .70), demonstrating that the Biotrainer provides similar information about activity. Subsequent calibration studies have demonstrated that the Biotrainer provides good predictive validity with adults and children. The intensity cut-point used to assess Biotrainer MVPA for the current study originated from a calibration study in children in which minutes of MVPA for the monitor were estimated from direct observation.
Procedures
Research assistants visited each school on a Monday and gave participants detailed instructions regarding how to record their activities using the PDPAR. Subjects were also given instructions regarding how to wear the Biotrainer and were encouraged to do so during all waking hours for 3 consecutive days (Tuesday, Wednesday, Thursday). Subjects completed the PDPAR for the same 3 days that the Biotrainer was worn. Logbooks were also kept to record the specific times the monitor was worn and reasons for time periods when the monitor was removed.

Data Processing

PDPAR
Data from the PDPAR were processed according to the guidelines established for this instrument. Each 30-minute period was converted into MET estimates based on the type and intensity of activities that were selected and in reference to the compendium of physical activities. If the type and intensity of a given activity did not appear congruent then the appropriate intensity for the type of activity listed was substituted. Periods that were ≥4 METS were classified as moderate to vigorous physical activity (MVPA). The total blocks of MVPA were tracked for the afternoon time frame (3:00 PM to 6:00 PM), the evening time frame (6:00 PM to 11:00 PM) and for the whole day (3:00 PM to 11:00 PM). Data were compiled separately by day and averaged across the 3 days to reflect typical activity levels.

Biotrainer Accelerometer
Accelerometer data were downloaded and processed using software provided by the manufacturer. Data were exported in 1-minute intervals and imported into SAS to evaluate compliance and to compute outcome data. Of primary importance in the current study was the accuracy of the prediction equation (internal validity). Therefore, stringent compliance standards were set to minimize measurement error in the objective accelerometer data. The screening procedures were the same as used in a previous study but more restrictive criteria were used to ensure that participants had clean data. Specifically, 30-min periods of consecutive 0 activity counts were flagged and identified as periods when the monitor may have been removed. If 4 or more periods (ie, 2+ hours) were identified on any given day, the participant was removed from all analyses. A total of 146 out of 240 participants met these compliance criteria (61%). An additional 25 participants were removed for missing anthropometric or PDPAR data, leaving 121 subjects in the final analysis.

The Biotrainer data were then processed to estimate minutes of MVPA performed each day. A child-specific activity cutpoint developed to capture the threshold for MVPA was used to process the data. Minutes with values above this point were classified as active minutes while minutes below were classified as nonactive. The total amount of MVPA was computed across different time intervals (afternoon and evening) to provide parallel estimates of activity levels as reported with the PDPAR. Data were compiled separately by day and averaged across the 3 days to reflect typical activity levels.

Data Analyses

Descriptive Analyses
Descriptive analyses were conducted to compare the activity levels of the 2 instruments. The analyses provided means and standard errors for PDPAR and Biotrainer data across the 3 days of monitoring. Data were segmented by gender and time period (afternoon and evening) to determine the consistency of the observed patterns of PA reported with the 2 instruments. A simple 2-way analysis of variance (ANOVA) was used to compare activity levels of males and females.

Calibration Analyses
A regression model was fit to determine the relationship between PDPAR bouts and Biotrainer minutes of PA. The model estimation method was restricted maximum likelihood with “school” included as a random effect to account for potential clustering within schools. A variety of models were tested to determine characteristics influencing fit. We evaluated models that included terms to allow for gender-specific y-intercepts and slopes (ie, separate linear equations for males and females). Body mass index (BMI) was also tested to determine if BMI accounted for additional variance in predicting PA levels. Finally, a quadratic term for the PDPAR variable was tested (PDPAR²) to determine if this provided a better fit for the data. To ensure a comprehensive analyses, we tested all combinations of models with the following predictors: PDPAR (blocks/day), PDPAR², BMI, gender (Males = 0, Females = 1), gender × PDPAR, and gender × PDPAR².

Significance testing was used to assess model fit. Main effect terms were included if significant at the P < .05 level, and interaction terms were included if significant at the P < .10 level. Neither BMI (P = .142), PDPAR² (P = .506), nor gender × PDPAR² (P = .771) was a significant predictor of objective PA, and so each was removed from the model. The main effect term for gender was also nonsignificant (P = .986), but this term was retained in the model to allow the gender × PDPAR interaction term (P = .058) to be included. As expected, the PDPAR was a strong predictor of accelerometer PA in all models (P < .001). The final model involved regressing Biotrainer MVPA minutes on PDPAR, gender, and gender × PDPAR using the regression model below.

\[
\text{MVPA} = \beta_0 + \beta_1 \times (\text{PDPAR}) + \beta_2 \times (\text{gender}) + \\
\beta_3 \times (\text{gender} \times \text{PDPAR}) + \text{error}
\]
All regression analyses were performed using PA data averaged across the 3 days to provide a more stable prediction equation for estimating agreement between the instruments.

**Validation Analyses**

The regression equation developed in the calibration analyses was then applied to a separate, holdout sample to evaluate the utility of the prediction algorithm. The developed equation was first evaluated by conducting a statistical test of differences between the measured and predicted values. Pearson Product moment correlations were then used to evaluate the overall strength of the association. Finally, Bland Altman plots were used to examine the distribution of error across the range of the activity estimates.

**Results**

A total of 91 participants were included in the calibration analysis and 30 participants in the validation analysis. Descriptive characteristics for the combined sample (calibration and validation group) are provided in Table 1. Average subject age (mean ± SE) was 12.3 ± 0.03 years, and average body weight was 49.5 ± 1.4 kg. The PDPAR calculations yielded a mean activity level of 3.2 ± 0.20 blocks/day while the estimated minutes of MVPA from the Biotrainer monitor was 28.6 ± 2.79 min/day. Plots comparing the PA levels segmented by time of day (afternoon and evening) and across the 3 days of monitoring are provided in Figure 1. No significant differences were seen in participants’ activity levels across the 3 days of monitoring for either the PDPAR (\( P = .348 \)) or the Biotrainer (\( P = .137 \)).

Mean PA levels for males and females are displayed by monitoring day (Tues-Thurs) for the PDPAR and Biotrainer monitor in Figure 2. When comparing Biotrainer minutes of MVPA by gender, males (42.1 ± 4.69) spent significantly more time in PA than females (17.1 ± 2.49) during the after-school time period (\( P < .001 \)). Similarly, PDPAR bouts of MVPA were greater for males (3.7 ± 0.26) than for females (2.7 ± 0.29) during the after-school time period (\( P = .016 \)). The final fitted calibration model included in the validation analysis is shown below.

\[
\text{Predicted MVPA} = 6.042 + 8.985 \text{PDPAR} - 0.1410 (\text{gender}) - 4.775 (\text{gender} \times \text{PDPAR})
\]

The above equation accounted for 36.6% of the variance in Biotrainer PA. To account for the differences between males and females in the model, gender was dummy-coded such that males = 0 and females = 1. This dummy-coding method produced a separate equation for each gender such that males had a slope of 8.985 and a y-intercept of 6.042, while females had a slope of 4.210 and a y-intercept of 5.901.

A cross-validation analysis on an independent holdout sample provided a comparison of the predicted and reported levels of PA. The prediction algorithm yielded an average of 23.9 ± 3.25 min/d of PA, while the Biotrainer data yielded estimates of 28.5 ± 6.18 min/d in the holdout sample. The difference between predicted minutes and Biotrainer minutes (4.6 ± 4.00) was not significant (\( t \)-value = 1.15, \( P = .261 \)). The 95% confidence interval for this difference in means was −3.6 min/d to 12.78 min/d. The correlation between the Biotrainer minutes of MVPA and the predicted minutes of MVPA in the holdout sample was high (\( r = .81 \)).

Bland Altman plots were also used to compare the accuracy of the predicted minutes across the range of PA levels (Figure 3). Examination of the plots revealed that there was more variation in predicted minutes of MVPA among participants (especially males) with greater activity levels. In general, it appears that the prediction equation may underestimate minutes of MVPA among male participants with relatively high levels of activity.

### Table 1 Descriptive Characteristics (Means and Standard Errors) for Total Sample and by Gender

<table>
<thead>
<tr>
<th></th>
<th>Total (n = 121)</th>
<th>Male (n = 56)</th>
<th>Female (n = 65)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>12.3 (0.03)</td>
<td>12.3 (0.06)</td>
<td>12.3 (0.04)</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.6 (0.01)</td>
<td>1.6 (0.02)</td>
<td>1.6 (0.01)</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>49.5 (1.39)</td>
<td>48.3 (2.57)</td>
<td>50.5 (1.37)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>19.6 (0.37)</td>
<td>19.4 (0.63)</td>
<td>19.8 (0.42)</td>
</tr>
<tr>
<td>PDPAR MVPA (a) (blocks/day)</td>
<td>3.2 (0.20)</td>
<td>3.7 (0.26)</td>
<td>2.7 (0.29)</td>
</tr>
<tr>
<td>Biotrainer MVPA (a) (min/day)</td>
<td>28.6 (2.79)</td>
<td>42.1 (4.69)</td>
<td>17.1 (2.49)</td>
</tr>
</tbody>
</table>

\(a\) Moderate plus vigorous physical activity.
Discussion

The current study compared PDPAR bouts of PA to objectively measured minutes of activity via accelerometry. Though the PDPAR has been validated in previous research, this is the first study that has sought to calibrate the instrument against objectively measured activity. Though the developers of the PDPAR never intended for it to be used to estimate time spent in PA, the ability to do so would clearly facilitate interpretation of research results. The current study provides new information about the PDPAR and the results may help to improve the utility of the PDPAR in future research.

This study yielded similar validity results as previous studies comparing the PDPAR to objective monitors. The correlation between average blocks and average minutes was 0.56 ($P < .001$) for the entire sample. Previous research with this sample using data from year 1 of the project yielded correlations of 0.56 to 0.65 between raw accelerometer counts and blocks. Pate, et al also reported moderate correlations between a 3-day physical activity recall (3DPAR) and total CSA accelerometer counts after both 7 days ($r = .51$) and 3 days ($r = .46$) of comparison. Trost et al showed similar correlations between PDPAR-based MET levels and accelerometer counts in 5th-grade children ($r = .51-.62$), though the association was weaker when comparing PDPAR blocks of MVPA ($>3$ METs) to accelerometer-based minutes of MVPA ($r = .19$).

Results were also similar for a study which employed a different self-report measure for PA. A study conducted by Cradock, et al compared 2 24-hour PA recalls to minutes of activity measured with TriTrac-R3D monitors (TTM) in 43 middle school students. Results showed
Figure 2 — Gender comparison of reported physical activity from the PDPAR (top panel) and recorded physical activity from the Biotrainer accelerometer (bottom panel) across 3 days of monitoring.

Figure 3 — Bland Altman plots comparing the accuracy of the prediction equations across the range of activity levels.
a correlation coefficient of 0.49 between the TTM and self-report MVPA (≥3 METS). Furthermore, the authors found that both male and female participants tended to overestimate (as a percent of day) time spent in MVPA relative to TTM (22.8% vs. 8.9%, \( P < .001 \)).

The modest (albeit significant) associations typically observed between self-report and objective monitors can be interpreted differently. Some may conclude that the self-report instruments are “valid” while others may cite the large variability in estimated PA to question the utility of self-report instruments. The current study attempts to understand the fundamental relationship between these 2 measures. Most researchers have assumed that youth overestimate activity levels due to cognitive problems in recall, social desirability, or because of inherently intermittent activity patterns.\(^1\)\(^,\)^\(^2\)\(^,\)^\(^27\)\(^,\)^\(^28\) However, it is also likely that much of the overestimation is due to the nature and wording of the assessments. Rather than trying to compare the raw output measures, the approach in this study was to develop a prediction algorithm that would allow data from the PDPAR to be equated to objective data from an accelerometer. Linking data to a common metric helps to facilitate comparisons across studies and may reduce confusion in the literature about reported levels of activity in youth. A better understanding of sources of bias for different segments of the population (eg, males vs. females or active vs. inactive) may help researchers comprehend sources of measurement error in self-reported PA. Measurement error models have been developed to improve the use of inherently complex dietary-recall data and similar work may help to improve the utility of self-reported PA.\(^29\)

The prediction equation from the current study suggests that, on average, youth obtain approximately 4 to 9 minutes of MVPA per PDPAR block, in addition to a baseline level of 6 minutes of MVPA when no blocks are reported. The higher coefficients for males suggest that males accumulate more minutes of MVPA (9.0 min/block) than females (4.2 min/block) for each reported PDPAR block. This difference in PA per reported block may be due to boys performing longer bouts of activity within each reported 30-min block of the PDPAR. It is also possible that girls perform PA less vigorously than boys, on average, which could result in a discrepancy between reported and measured activity levels if the less-intense activity falls below the accelerometer cut-point for MVPA.

The utility of the prediction equations were evaluated by applying the equations to an independent, holdout sample. We found a nonsignificant difference between measured and predicted levels of PA \( (P = .261) \). This difference \( (4.6 \text{ min/d of MVPA}) \) was relatively small considering the somewhat insensitive metric obtained from the PDPAR (30-minute blocks). The present results demonstrate promise for the utilization of a simple calibration equation to improve the utility of the PDPAR. More robust statistical techniques and sampling procedures may yield more robust calibrations but the purpose of the current study was to evaluate the utility of this relatively simple calibration procedure.

This study has several limitations which are important to note. Due to the data collection method and the local demographics (ie, 9 schools in the Midwest) participants from the current sample were similar in age and primarily Caucasian. Future research should assess the relationship between PDPAR bouts and objectively measured minutes of activity among different age groups and ethnicities to account for potential self-report differences among these demographic groups. The relationship between PDPAR bouts and accelerometer minutes also may be influenced by the type of accelerometer and the choice of monitor cut-point used to define MVPA. Numerous studies have demonstrated that accelerometer MVPA outcomes are highly dependent on the choice of intensity cut-point.\(^30\)\(^,\)^\(^31\) One way to improve the predictive accuracy of this type of calibration procedure may be to use a triangulation approach by employing multiple PA measures. This would take into account the limitations of the various cut-points and yield associations that are closer to truth. Another limitation was that the validation analyses used a relatively small hold-out sample \( (n = 30) \) which restricted our ability to find significant differences between the PA estimates. However, the relatively strong correlation \( (r = .81) \) between predicted and objective minutes of MVPA helps to provide support for the validity of the equation. A final limitation is that the proposed conversion algorithm is that it is based on group data and is, therefore, not valid for estimating minutes of MVPA from individual PDPAR data. This limitation is not unique to the current study since most accelerometer calibration equations developed to date are only suitable for group estimates. Additional research on the calibration of self-report instruments with objective PA data may help to improve the utility of field-based research on PA.

References