A Physical Activity Screening Measure for Use With Adolescents in Primary Care

Judith J. Prochaska, MS; James F. Sallis, PhD; Barbara Long, MD, MPH

Objective: To develop a reliable and valid physical activity screening measure for use with adolescents in primary care settings.

Study Design: We conducted 2 studies to evaluate the test-retest reliability and concurrent validity of 6 single-item and 3 composite measures of physical activity. Modifications were based on the findings of the 2 studies, and a best measure was evaluated in study 3. Accelerometer data served as the criterion standard for tests of validity.

Results: In study 1 (N=250; mean age, 15 years; 56% female; 36% white), reports on the composite measures were most reliable. In study 2 (N=57; mean age, 14 years; 65% female; 37% white), 6 of the 9 screening measures correlated significantly with accelerometer data. Instead, we developed a single measure assessing accumulation of 60 minutes of moderate to vigorous physical activity. Evaluated in study 3 (N=148; mean age, 12 years; 65% female; 27% white), the measure was reliable (intraclass correlation, 0.77) and correlated significantly ($r=0.40, P<.001$) with accelerometer data. Correct classification (63%), sensitivity (71%), and false-positive rates (40%) were reasonable.

Conclusion: The “moderate to vigorous physical activity” screening measure is recommended for clinical practice with adolescents.

Arch Pediatr Adolesc Med. 2001;155:554-559

Research supports the benefits of physical activity in young people, for health both in adolescence and later in adulthood. Healthy People 2010 recommends vigorous physical activity (VPA) for at least 20 minutes at a time, 3 times per week, and accumulation of at least 30 minutes per day of moderate physical activity (MPA) most days of the week. Guidelines specifically developed for youth recommend accumulation of 60 minutes of moderate or greater intensity activity on most days of the week.

National survey data of adolescents in grades 9 through 12 indicate that 65% meet the vigorous and 27% meet the 30-minute moderate physical activity guidelines. Girls, ethnic minorities, and older youth are less likely to meet the recommendations. National prevalence data, however, are based on self-report measures of questionable validity. When more objective measures are used (eg, heart rate monitor and electronic monitor), estimates of the proportion of youth meeting the guidelines drop dramatically. At this time, no data are available for the 60-minute MPA guideline. Most young people could benefit from increasing their participation in physical activity.

The American Academy of Family Physicians, the American Academy of Pediatrics, the American College of Sports Medicine, and the American Medical Association recommend physical activity counseling for children and adolescents. Healthy People 2010 objective 1-3 is to increase the proportion of individuals appropriately counseled about health behaviors. Most adolescents (70%) visit a physician at least once a year, and for health information, adolescents report relying most on parents and physicians. It is estimated that less than one third of patients aged 6 to 17 years receive counseling on physical activity from their primary care providers.

To guide physical activity counseling, accurate and reproducible screening measures are needed. For the clinical setting, the measures must also be brief enough to be practical, assess targeted behaviors, and yield clinically useful scores. The purpose is not to comprehensively assess individuals’ physical activity habits,
but rather to identify individuals not meeting the guidelines who could benefit from counseling. In research and clinical practice, self-report has typically been the method of choice with older children. However, there are no validated self-report measures for youth that are brief and specific enough for use in primary care. This article describes the process of developing a reliable and valid physical activity screening measure for use with adolescents in primary care settings. We present data from 3 separate studies.

**STUDY 1**

In study 1, we examined test-retest reliability of 9 measures of physical activity. Test-retest reliability indicates temporal stability, or how constant scores remain from one testing occasion to another. A measure with minimal random variation is desired. Physical activity behaviors are, however, expected to demonstrate some natural variation over time.

**SUBJECTS AND METHODS**

**Study Participants**

Subjects were recruited from required classes in 2 high schools and 2 middle schools in San Diego, Calif, and Pittsburgh, Pa. The study protocol received approval for use of human subjects. Subjects had to provide assent, obtain passive parental consent (1 school required written parental consent), and speak English. Participation in the study was 79%, with lower participation at the school requiring active (55%) vs passive (90%) parental consent. Of 278 subjects, 250 completed assessments at both time points. Mean age of the sample was 14.6 years (SD, 1.4 years); 56% were female. Ethnic distribution was white (36%), Asian/Pacific Islander (25%), African American (17%), Hispanic (9%), and other (13%).

**Measures**

Vigorous physical activity measures corresponded with national guidelines and were based on items from the Youth Risk Behavior Survey. Two single-item measures assessed the number of days individuals had engaged in bouts of VPA for at least 20 minutes at a time during the past 7 days and for a typical week. Vigorous physical activity was defined as “usually makes you breathe hard or feel tired most of the time” and examples were provided: jogging, soccer, and “aggressive” skateboarding. The 2 items were also averaged to form a composite measure. The measures yielded a score of days per week the adolescent accumulated the specified minutes of MPA. Five or more days per week met the guidelines. The demographic measure assessed age, sex, and ethnicity.

**Procedure**

The measures were initially piloted with a small sample of adolescents (n=6), diverse in age, reading level, and ethnicity. In study 1, subjects completed the measures twice at an approximate interval of 2 weeks. Students completed the surveys at school, supervised by research staff.

**Statistical Methods**

We analyzed data for the full sample and for boys and girls within younger (grades 7-8) and older (grades 9-12) age groups separately. A multivariate general linear model tested differences on the measures by age, sex, and race. One-way model intraclass correlation coefficients (ICCs) evaluated reliability at the item level. We also computed κ statistics to evaluate the measures’ reliability for classifying subjects as meeting or not meeting guidelines. Landis and Koch interpret values of κ as follows: less than 0%, poor; 0% to 20%, slight; 21% to 40%, fair; 41% to 60%, moderate; 61% to 80%, substantial; and 81% to 100%, almost perfect.

**RESULTS**

**Descriptive Statistics**

Physical activity scores for the full sample are summarized in Table 1. Scores were similar for typical-week and past-7-days reference periods. Reports were much lower on the 60-minute than the 30-minute MPA measure. There were significant differences in reports by sex (F6,233=4.93; P=.001) and race (F21,623=2.37; P=.001) but not age. On all 9 measures, boys reported significantly more physical activity than girls, and white and Asian/Pacific Islander students reported greater physical activity than students of other races.

**Test-Retest Reliabilities**

There was no consistent trend in the strength of reliabilities when analyzed by age or sex. Reliability statistics for the full sample are summarized in Table 1. Reliability was strengthened with the composite measures. The VPA (ICC, 0.76) and 60-minute MPA (ICC, 0.79) composite measures had the strongest reliabilities. The κ statistics ranged from 45% to 61%. Only the 60-minute MPA composite reached the criterion to be considered substantial (κ=61%).

**STUDY 2**

In study 2, we examined concurrent validity of the measures. The greatest obstacle to validating physical activity
accelerometer has been validated for youth\textsuperscript{21,22} and is a Computer Science and Applications (CSA, Shalimar, Fla) cal activity can now be compared with self-reports. The levels. Detailed data on frequency and intensity of physi-

assessments has been the lack of an adequate criterion standard.\textsuperscript{20} Recently developed electronic accelerometers offer the advantage of storing minute-by-minute activity levels. Correlations, we explored use of the screening measures for identifying individuals not meeting physical activity guidelines.

**SUBJECTS AND METHODS**

**Study Participants**

Study 2 was conducted in the San Diego schools involved in study 1. Half of the sample had participated in study 1. Active consent and assent were required. Of 62 subjects, 57 had sufficient CSA data to be included in analyses. Mean age of the sample was 13.9 years (SD, 1.7 years); 37 (65\%) were female. Ethnic distribution was white subjects (21 subjects [37\%]), Asian/Pacific Islander (14 [25\%]), Hispanic (7 [12\%]), African American (2 [4\%]), and other (13 [23\%]). (Percentages have been rounded and may not total 100.)

**Measures**

Content, administration, and scoring of the physical activity and demographic measures were identical to those described in study 1.

The CSA activity monitor (model 7164) is a small (5.1 × 3.8 × 1.5 cm), durable, lightweight (45 g), uniaxial accelerometer measuring integrated accelerations in the vertical plane. The CSA has been shown to be a valid tool for quantifying children's activity levels in laboratory and field settings; correlations with heart rate monitoring range from 0.50 to 0.74.\textsuperscript{21,22} Limitations of the CSA monitor include the following: (1) it is uniaxial and thus underestimates activities that produce little vertical trunk movement (eg, bicycling), and (2) it is not waterproof and cannot assess activities performed in the water. This study used the CSA's summed magnitude mode, considered to reflect the duration, frequency, and intensity of activity. Assessments were made at a 1-minute sampling interval.

Subjects wore a CSA monitor on their right hip, secured to an elastic belt, for 7 days. Subjects were instructed to wear the monitor at all times, except when sleeping, showering, or swimming. During the assessment period, research staff called subjects to record any physical activities not well assessed by the monitor. At the end of the week, subjects completed the self-report measures.

**Statistical Methods**

The CSA data were downloaded to a personal computer. A Q-basic software program developed by Trost and colleagues\textsuperscript{2} calculated total minutes per day spent in MPA and VPA. Physical activity intensity was defined as 3.00 to 5.99 METs for moderate and 6 METs or more for vigorous, where 1 MET is the metabolic equivalent of an individual sitting at rest. The program is calibrated on the basis of laboratory studies of oxygen consumption during treadmill locomotion with youth 6 to 17 years old.\textsuperscript{23} The equation adjusts for subject's age and sex. The program also calculates the number of 20-minute bouts of VPA.

We considered a day with less than 8 hours of recorded activity as missing and required 5 days of CSA data for analyses. We calculated average minutes of VPA and MPA per day. We also created adjusted CSA variables based on subjects’ reports of time spent in activities not well assessed by the CSA (eg, bicy-

cling) or done while the accelerometer was not worn (eg, swimming). For analysis of correct classification, we calculated number of days subjects engaged in 20-minute bouts of VPA and accumulated 30 minutes and 60 minutes of MPA. The CSA data were entered into SPSS version 8.0 statistical software (SPSS Inc, Chicago, Ill).\textsuperscript{24} We evaluated validity in 3 steps. First, Pearson correlations tested the association between self-report and accelerometer data. Second, we directly compared self-report values with CSA data. Last, we calculated classification rates for measures with the strongest validity coefficients. Subjects were coded as meeting or not meeting guidelines on the basis of CSA and self-report data. For VPA, we had to use a less stringent criterion than the traditional guideline because so few sub-

<table>
<thead>
<tr>
<th><strong>Table 1. Descriptive Statistics and Test-Retest Reliability for Physical Activity Measures\textsuperscript{*}</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intensity</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Vigorous</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Moderate</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Moderate</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{*} N = 250. ICC indicates intraclass correlation coefficient.
jeks engaged in extended bouts of activity at this intensity. A guideline of accumulating 60 minutes or more of VPA for the week was chosen. For the self-report measures, cutoff points were 3 or more days a week for VPA and 5 or more days a week for MPA. We calculated the correct classification rate as a proportion of agreement between the self-report measure and the CSA for classifying subjects with respect to the guidelines. We calculated sensitivity as the proportion of subjects not meeting the guideline on the basis of CSA data but identified as not meeting the guideline by the self-report measure.

RESULTS

Validity Correlations

Correlations were significant for the VPA and the 60-minute MPA measures, but not for the 30-minute MPA measures (Table 2). Correlations ranged from 0.20 to 0.46 and were strongest for the composite measures. Scatterplots did not reveal correlations to be obviously affected by outliers. Adjusting CSA data with subject-composite measures with the strongest validity data were evaluated for correct classification rates. For the 60-minute MPA composite, the correct classification rate (78%) and sensitivity (80%) were good; the false-positive rate was 4%. Sensitivity (38%) and the correct classification rate (58%) were low for the VPA composite because of problems with overreporting. The false-positive rate was 0%. A potential alternative could be to assess general physical activity (ie, moderate to vigorous physical activity [MVPA]) instead of VPA and MPA separately. Accumulation of 60 minutes of MVPA corresponds with recent recommendations for youth physical activity. We correlated the 60-minute MPA composite with minutes of MVPA assessed by the CSA monitor. Correlation and classification data were not as strong as with CSA minutes of MPA, but the relationship was significant (r=0.38; P=0.003). Correct classification rate (70%) and sensitivity (83%) were reasonable; the false-positive rate (67%) was high.

Descriptive Statistics

The sample averaged 11 minutes of VPA and 71 minutes of MPA on the basis of CSA data. Standard deviations were large, in many cases larger than the mean values, indicating great variability among subjects. Very few subjects met the VPA guideline of 20-minute bouts (only 3 or 57 subjects). If the guideline allowed for accumulation of VPA, 35% of subjects (20/57) would meet the guideline of 60 minutes or more per week. Five subjects (9%) met the 60-minute MPA guideline. In contrast, 45 (79%) and 14 (25%) of the sample self-reported meeting guidelines for VPA and MPA, respectively. Descriptive statistics showed notable differences between self-report and CSA data. For the VPA composite measure, subjects overreported participation a mean of 3.3 days per week. Subjects overreported participation in MPA, but the differences were less—a mean of 1.4 days per week for both MPA composite measures.

Classification Rates

Measures with the strongest validity data were evaluated for correct classification rates. For the 60-minute MPA composite, the correct classification rate (78%) and sensitivity (80%) were good; the false-positive rate was 4%. Sensitivity (38%) and the correct classification rate (58%) were low for the VPA composite because of problems with overreporting. The false-positive rate was 0%. A potential alternative could be to assess general physical activity (ie, moderate to vigorous physical activity [MVPA]) instead of VPA and MPA separately. Accumulation of 60 minutes of MVPA corre-

<table>
<thead>
<tr>
<th>Self-report Measure</th>
<th>Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>VPA</td>
<td></td>
</tr>
<tr>
<td>Typical week</td>
<td>0.31†</td>
</tr>
<tr>
<td>Past 7 d</td>
<td>0.36‡</td>
</tr>
<tr>
<td>Composite</td>
<td>0.37‡</td>
</tr>
<tr>
<td>30-min MPA</td>
<td></td>
</tr>
<tr>
<td>Typical week</td>
<td>0.20</td>
</tr>
<tr>
<td>Past 7 d</td>
<td>0.26</td>
</tr>
<tr>
<td>Composite</td>
<td>0.26</td>
</tr>
<tr>
<td>60-min MPA</td>
<td></td>
</tr>
<tr>
<td>Typical week</td>
<td>0.46‡</td>
</tr>
<tr>
<td>Past 7 d</td>
<td>0.37‡</td>
</tr>
<tr>
<td>Composite</td>
<td>0.46‡</td>
</tr>
</tbody>
</table>

* N = 57. CSA indicates Computer Science and Applications; VPA, vigorous physical activity; and MPA, moderate physical activity. †P < .05. ‡P < .01.

STUDY 3

On the basis of the findings from studies 1 and 2, we modified our assessment strategy to better match the types and patterns of youth physical activity. In study 2, the 60-minute MPA composite out-performed all other measures and significantly correlated with minutes of MVPA assessed by the CSA monitor. Correlation and classification data were not as strong as with CSA minutes of MPA, but the relationship was significant (r=0.38; P=0.003). Correct classification rate (70%) and sensitivity (83%) were reasonable; the false-positive rate (67%) was high.

SUBJECTS AND METHODS

Study Participants

Study 3 was conducted in the San Diego middle school involved in studies 1 and 2. Two years had passed, so none of the students had participated in the previous studies. Active consent and assent were required. The study sample consisted of 138 subjects (65% female) with a mean age of 12.1 years (SD, 0.9 year). Ethnic distribution was white (27%), Asian/Pacific Islander (24%), Hispanic (3%), African American (7%), multiracial (23%), and other (14%).
Measures

The MVPA measure assessed the number of days subjects had accumulated 60 minutes of MVPA during the past 7 days and for a typical week. The measure defined physical activity broadly as “increases your heart rate and makes you get out of breath some of the time” and did not specify intensity. A composite average of the 2 items yielded a score of days per week the adolescent accumulated 60 minutes of MVPA. Five or more days per week met the guideline.

The CSA monitor was the comparison measure for assessing concurrent validity.

Procedures

Subjects wore the CSA on their right hip on a standard belt provided for the study. Subjects wore the CSA every day, all day, for a 5-day period. The CSA monitoring occurred the week before completion of the MVPA composite. Subjects completed the paper-based survey individually, in a group setting, supervised by research staff. Subjects completed the same measure on a computer either the same day or up to 1 month after completion of the paper version.

Statistical Methods

We evaluated reliability between the paper-based survey and the retest on the computer. We examined differences in reliability by time to retest both by adding time as a covariate to ICC analyses and by running analyses separately for 4 retest subgroups: same day, 24-hour retest, 1 to 6 days, and 1 week or more. We also examined reliability of the measure for classifying subjects with respect to physical activity guidelines. We calculated χ statistics for the full sample and for subgroups based on time to retest.

We analyzed CSA data by means of the Q-basic program to yield minutes of MVPA accumulated in a day. We considered days with less than 8 hours of recorded activity as missing and required a minimum of 5 days of data for validity analyses. We correlated CSA mean minutes of MVPA with the 60-minute MVPA composite. We calculated sex-specific correlations. Since the CSA assessment did not cover a full week, it was not possible to determine whether subjects were meeting the guideline of 5 or more days per week. For correct classification rates, we used mean minutes of MVPA per day with a cutoff point of 60 minutes.

RESULTS

Descriptive Statistics

Scores on the 60-minute MVPA composite ranged from 0 to 7, with a mean of 4.8 (SD, 2.0) days per week. More than half (53%) of the sample reported meeting the guideline of 5 or more days per week. Boys (mean, 5.2; SD, 1.8) reported engaging in more physical activity than girls (mean, 4.5; SD, 2.0) (F1,137 = 3.96; P = .049). The measure correlated negatively with age (r = −0.27; P = .007).

Ninety-nine subjects had 5 or more days of CSA data. The sample averaged 85 minutes (SD, 30 minutes) of MVPA per day (range, 11-162 minutes). Boys (mean, 100 minutes; SD, 30 minutes) were more active than girls (mean, 77; SD, 27) (F1,97 = 16.36; P < .001). Age negatively correlated with CSA data (r = −0.39; P < .001). A majority of the sample (79%) averaged at least 60 minutes per day of MVPA.

Reliability

For the full sample, the ICC was 0.77. With time to retest as a covariate, the ICC was 0.76. Reliabilities ranged from ICC = 0.88 for a same-day retest (n = 42) to ICC = 0.53 for a retest at up to 1 month (n = 31). The overall χ statistic (61%) was substantial. The χ values ranged from 84% for a same-day retest to 36% for a retest at up to 1 month.

Validity

The 60-minute MVPA composite correlated significantly with CSA data (r = .40; P < .001). The association was stronger for boys (r = .42; P = .1; n = 36) than for girls (r = .32; P = .01; n = 63). There were no problems with outliers. Intraclass correlation was 0.77 (n = 138; χ = 61%). Correct classification rate for the full sample was 63%, with 71% sensitivity, and a 40% false-positive rate.

We conducted 3 studies with the objective of creating a reliable and valid measure of adolescent physical activity. In studies 1 and 2, we evaluated 9 measures. Subjects provided reliable reports on the measures but had difficulty estimating participation in continuous bouts of activity and distinguishing between intensity levels. While accelerometers may not capture all activity performed, even when self-reported log data were used, errors in reporting remained large. Findings from the current study are consistent with those of other objective monitoring studies that have found that few young people engage in continuous 20-minute bouts of physical activity.20 From these findings, we created a single measure assessing accumulation of MVPA. The measure is consistent with recent recommendations for youth to accumulate 60 minutes of MVPA on most days of the week.36 Study 3 evaluated reliability and validity of this measure.

Reliability (ICC = 0.77) and validity (r = 0.40) of the 60-minute MVPA measure were comparable to those reported in the literature. In a recent review of 17 self-report instruments for youth physical activity, reliabilities ranged from 0.60 to 0.98, with stronger reliability observed for same-day retests.13 Validity correlations ranged widely from 0.02 to 0.88; only 2 measures had correlations above 0.50.13 Correlations provide an indication of relative validity. Few studies have examined validity of a screening measure for correctly classifying subjects. The correct classification rate (63%), sensitivity (71%), and false-positive rate (40%) of the 60-minute MVPA measure were reasonable. Physical activity counseling is low risk and can potentially benefit all. For clinical screening, we considered sensitivity more important.
PACE+ Adolescent Physical Activity Measure

**Physical activity** is any activity that increases your heart rate and makes you get out of breath some of the time.

**Physical activity** can be done in sports, playing with friends, or walking to school.

Some examples of **physical activity** are running, brisk walking, rollerblading, biking, dancing, skateboarding, swimming, soccer, basketball, football, and surfing.

Add up all the time you spend in physical activity each day (don’t include your physical education or gym class).

**P1** Over the past 7 days, on how many days were you physically active for a total of at least 60 min per day?

- □ 0 days
- □ 1 days
- □ 2 days
- □ 3 days
- □ 4 days
- □ 5 days
- □ 6 days
- □ 7 days

**P2** Over a typical or usual week, on how many days are you physically active for a total of at least 60 min per day?

- □ 0 days
- □ 1 days
- □ 2 days
- □ 3 days
- □ 4 days
- □ 5 days
- □ 6 days
- □ 7 days

Scoring: (P1 + P2)/2 < 5 indicates not meeting physical activity guidelines.

Sixty-minute screening measure for moderate to vigorous physical activity: PACE+ (Patient-Centered Assessment and Counseling for Exercise Plus Nutrition).

than the false-positive rate. The 60-minute MVPA composite is a reasonable method for assessing participation in overall physical activity and for assessing achievement of current guidelines.

We conducted 3 studies to evaluate the reliability and validity of multiple self-report measures of youth physical activity. The sample in study 1 was large and drawn from 2 geographically distinct areas. Samples in all 3 studies were ethnically diverse and represented different developmental levels. The findings support use of the measures with youth diverse with respect to age, sex, and race. The validity studies used an objective physical activity comparison measure, and measures were evaluated for screening individuals in relation to clinically relevant health guidelines. These are among the few studies of youth self-report measures to validate reports of absolute amount of physical activity.13

A final single measure, the 60-minute MVPA screening measure, is recommended for clinical practice (Figure). The screening measure has been incorporated into the PACE+ computer-mediated physical activity program for adolescents in primary care.28 The measure is brief and easy to score, and it yields clinically meaningful scores. The measure provides a reliable estimate of adolescents’ physical activity behavior and correlates significantly with an objective measure of physical activity. As this article demonstrates, measurement development is an iterative process, with measures being modified, evaluated, and refined.

Accepted for publication November 3, 2000.

Studies 1 and 2 were supported by a student oncology grant from the American Cancer Society California Division, Oakland, Calif. Study 3 was supported by a predoctoral psychosocial fellowship from the American Cancer Society California Division, Oakland, and a dissertation grant from the American College of Sports Medicine, Indianapolis, Ind.

We thank Rene Carreno, Corina Fisher, Diane Wade, Miki Watanabe, David Cohen, and Béatrice Schmid, MA, for help with data collection and data entry for the 3 studies.

Presented at the annual meeting of the American College of Sports Medicine, Seattle, Wash, June 4, 1999.

Corresponding author: Judith J. Prochaska, MS, San Diego State University, 6363 Alvarado Ct, Suite 250, San Diego, CA 92120 (e-mail: prochask@sunstroke.sdsu.edu).

**REFERENCES**


(REPRINTED) ARCH PEDIATR ADOLESC MED/VOL 155, MAY 2001 WWW.ARCHPEDIATRICS.COM 399

©2001 American Medical Association. All rights reserved.