Establishing and Evaluating Wrist Cutpoints for the GENEActiv Accelerometer in Youth

CHRISTINE A. SCHAEFER[1], CLAUDIO R. NIGG[2], JAMES O. HILL[3], LOIS A. BRINK[3], and RAYMOND C. BROWNING[1]

[1]Colorado State University, Fort Collins, CO; [2]University of Hawaii, Honolulu, HI; and [3]University of Colorado Denver, Aurora, CO

ABSTRACT

SCHAEFER, C. A., C. R. NIGG, J. O. HILL, L. A. BRINK, and R. C. BROWNING. Establishing and Evaluating Wrist Cutpoints for the GENEActiv Accelerometer in Youth. *Med. Sci. Sports Exerc.*, Vol. 46, No. 4, pp. 826–833, 2014. **Purpose**: This study aimed to establish physical activity (PA) intensity cutpoints for a wrist-mounted GENEActiv accelerometer (ACC) in elementary school-age children. A second purpose was to apply cutpoints to a free-living sample and examine the duration of PA based on continuous 1-s epochs. **Methods**: Metabolic and ACC data were collected during nine typical activities in 24 children age 6–11 yr. Measured $\dot{V}O_2$ values were divided by Schofield-estimated resting values to determine METs. ACC data were collected at 75 Hz, band pass filtered, and averaged over each 1-s interval. Receiver operator characteristic curves were used to establish cutpoints at sedentary (≤1.5 METs), light (1.6–2.99 METs), moderate (3.0–5.99 METs), and vigorous (≥6 METs) activities. These cutpoints were applied to a free-living independent data set to quantify the amount of moderate–vigorous PA (MVPA) and to examine how bout length (1, 2, 3, 5, 10, 15, and 60 s) affected the accumulation of MVPA. **Results**: Receiver operator characteristic yielded areas under the curve of 0.956, 0.946, and 0.940 for sedentary, moderate, and vigorous intensities, respectively. Cutpoints for sedentary, moderate, and vigorous intensities were 0.190*g*, 0.314*g*, and 0.998*g*, respectively. Intensity classification accuracies ranged from 27.6% (light) to 88.7% (vigorous) when cutpoints were applied to the calibration data. When applied to free-living data (n = 47 children age 6–11 yr), estimated daily MVPA was 308 min and decreased to 14.3 min when only including 1-min periods of continuous MVPA. **Conclusions**: Cutpoints that quantify movements associated with moderate–vigorous intensity, when applied to a laboratory protocol, result in large amounts of accumulated MVPA using the 1-s epoch compared to prior studies, highlighting the need for representative calibration activities and free-living validation of cutpoints and epoch length selection. **Key Words:** PHYSICAL ACTIVITY, ACCELEROMETRY, YOUTH, CALIBRATION, VALIDATION

Accurate, objective physical activity (PA) monitoring is crucial to our understanding of current activity levels as well as for evaluating the effectiveness of interventions aiming to increase PA. Accelerometers (ACCs) are the most widely used objective measure of PA in both children and adults (25). Multiple ACC devices are now commercially available. Historically, the device software has applied proprietary processing algorithms to the unfiltered acceleration signal. This method results in count values generated by the devices that are difficult or, in many cases, impossible to compare across devices. This significant limitation has prompted the PA research community to support and encourage the development of devices that collect and store raw (i.e., high-frequency, preprocessed, unfiltered) ACC data (14). Fortunately, advances in data storage and battery life have made these devices readily available to the researcher (12). This new generation of ACC devices should facilitate comparisons across studies and devices (e.g., GENEActiv and ActiGraph GT3X+), permit robust PA quantification (including activity classification) and improve estimates of activity intensity (7,14). However, particularly in the short term, activity classification approaches are unlikely to take the place of intensity cutpoints. While recent research has investigated the accuracy of activity classifiers in children (33), other studies suggest that, although these classifiers are accurate when applied to data collected in a laboratory, classification accuracy is relatively poor when applied to free-living data (8). Given that PA guidelines continue to be recommended in minutes of moderate–vigorous PA (MVPA), and no robust activity classification system has been validated for use in children, researchers still require a way to quantify minutes of MVPA. Therefore, cutpoints must be established and evaluated for ACC devices that collect and store raw data, particularly to understand how interventions affect children's accumulation of daily MVPA.

Address for correspondence: Christine A. Schaefer, M.S., 220 Moby B Complex, Colorado State University, Fort Collins, CO 80523-1582; E-mail: Christine.a.schaefer@gmail.com.
Submitted for publication October 2012.
Accepted for publication August 2013.
Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.acsm-msse.org).

0195-9131/14/4604-0826/0
MEDICINE & SCIENCE IN SPORTS & EXERCISE®

DOI: 10.1249/MSS.0000000000000150

Although multiple investigators have cited the need to collect raw ACC data to assess PA, few have done so (7,14). Given the relatively recent nature of the capability to collect raw ACC data in multiday, free-living studies, we do not yet have clear data processing guidelines, and there are unanswered questions regarding interpretation of the data. For example, over what period of time (epoch) should researchers process data collected at frequencies of up to 100 Hz? While a shorter epoch allows improved temporal resolution of PA, it could be argued that, at very short epochs (<1 s), the device is quantifying movement that may not equate to purposeful (e.g., hand movement from reading, computer gaming, etc.) PA. However, direct observation of children indicates that they engage in short, intermittent periods of movement, with higher-intensity activities lasting an average of 3 s (4). This suggests that epochs of 1–2 s may be necessary to provide the resolution needed to detect all PA in children.

Another important consideration in quantifying PA accumulation is the duration of the activity bouts taking place. Here, we define the bout as a continuous period of PA above a given intensity threshold. Current guidelines for adults state that MVPA be accumulated in a minimum of 10 consecutive minutes (23). To date, no such recommendation has been established for children, but children are recommended to participate in at least 60 min of MVPA most days of the week. However, as noted above, it is well known that children typically do not choose to participate in long periods of sustained PA but rather engage in sporadic movements (4,6,18). In one of the only studies to examine the effects of minimum bout duration on reported PA accumulation in children, the number of moderate bouts of activity decreased from an average of ~193 bouts per day using a 2-s minimum bout duration (the epoch length of this study) to 5.3 bouts using a 20-s bout minimum (5). This highlights the significance of bout duration in the interpretation of daily PA accumulation in children. Importantly, no studies have yet reported PA accumulation using a 1-s epoch or bout in children.

The move toward collecting raw ACC data is logical and necessary. While this raw ACC data will eventually allow a much more detailed understanding of PA, estimating minutes of MVPA remains an important objective. One of the devices currently capable of raw ACC data collection is the GENEActiv ACC (Activinsights Limited, Cambridge, UK). It is waterproof and has been validated for wrist placement (10). The wrist is an attractive location, particularly in children, given improvements in compliance typically observed. For example, in our large Intervention of PhysicaL Activity in Youth (IPLAY) study of approximately 1400 elementary school-aged children older than 3 yr, we have achieved a compliance rate of ~92%–97% (34) (see Methods for additional IPLAY details). However, to date, only one study has attempted to create cutpoints specific to the GENEActiv device when placed on the wrist in children. Therefore, the primary purpose of this study was to establish wrist-based cutpoints for the GENEActiv ACC in children ages 6–11 yr. We hypothesized that the GENEActiv would accurately discriminate between sedentary, light, moderate, and vigorous activities. Our secondary aim was to apply these cutpoints to a free-living sample and to examine how the estimated accumulation of minutes of MVPA is affected by the bout length criteria. We hypothesized that children's PA accumulation would decrease significantly as bout length increased.

METHODS

We conducted a calibration experiment on 24 children ages 6–11 yr (Table 1). We placed the GENEActiv ACC on children's nondominant wrist while they participated in nine activities in a laboratory. Approval for this study was provided by the Institutional Review Board for Human Subjects Research at Colorado State University. All children and parents signed informed assent and consent forms, respectively, before children's participation in the study.

Study Design and Activities

Before participation, we conducted a phone screening with the parent to assess any contraindications to exercise. We asked that children arrive at the Physical Activity Laboratory having fasted for a minimum of 2 h. Participants typically came in pairs, which allowed them to feel more comfortable in the laboratory setting. Upon arrival, staff explained the study details to the child and obtained informed assent from the child and consent from the parent. Next, we measured each child's weight (Health-o-Meter Professional, Model 349KLX) and height (Detecto, Webb City, MO). We then fitted one of the children with the portable indirect calorimetry system as well as the GENEActiv device. Upon completion of the activities by the first child,

TABLE 1. Subject characteristics for calibration sample and independent IPLAY subsample.

	Subjects (*n*)	Height (cm)	Weight (kg)	Age (yr)
Calibration/validation study				
Girls	13	140.0 ± 7.3	34.0 ± 5.8	9.5 ± 1.1
Boys	11	141.4 ± 12.5	35.2 ± 10.1	9.3 ± 1.3
Total	24	140.6 ± 9.8	34.6 ± 7.9	9.4 ± 1.2
IPLAY subsample				
Girls	20	139.65 ± 11.1	36.9 ± 11.1	9.6 ± 1.5
Boys	27	136.7 ± 14.1	37.9 ± 14.9	8.9 ± 1.9
Total	47	138.0 ± 12.9	37.5 ± 13.3	9.2 ± 1.8

Values are mean ± SD.

study staff recalibrated the indirect calorimeter for the second child, who then completed the nine activities sequentially. The protocol began with an initial 6-min resting trial, during which children were asked to lie quietly in a clinical bed while watching a parent-approved DVD. Additional activities included (in order) coloring (seated), Lego building (seated on the floor), Wii Sports Tennis, Wii Sports Boxing, treadmill walking at two speeds (0.75 and 1.25 m·s^{-1}), jogging (1.75 m·s^{-1}), and running (2.25 m·s^{-1}). Each activity trial lasted 6 min. To synchronize the metabolic system with the ACC data, we simultaneously placed markers in the metabolic data file and on the ACC device, marking the end of each trial. We then analyzed the last 2 min of metabolic and accelerometry data preceding the event markers. On average, the study visit lasted 1.5 h per child (~3 h per pair of children).

Instrumentation

Accelerometry. The GENEActiv ACC is lightweight (16 g), triaxial, and waterproof. It collects raw acceleration data (range = ±8*g*). It has storage capabilities of 0.5 Gb at recording frequencies ranging from 10 to 100 Hz and can collect data for up to 7 d at 100 Hz. Data are downloaded using a USB 2.0 Charging Cradle. Devices were calibrated by the manufacturer before use. We collected data at 75 Hz and downloaded the data using the GENEActiv software (Version 2.1). We used a customized MATLAB program (MATLAB v 12.0; MathWorks, Natick, MA) to filter the data (band pass, nonrecursive with cutoff frequencies of 0.2 and 15 Hz). We filtered the data to remove gravitational acceleration and reduce the inclusion of accelerations associated with the device moving relative to the wrist. Although we did not perform a frequency analysis of the ACC in a "noise-free" protocol, studies have reported that the frequency content of ground reaction forces (most relevant to acceleration) during human locomotion are <9–17 Hz (1,16). We then calculated an average gravity-subtracted signal vector magnitude (SVM_g) for each second (see Equation 1, f = sampling frequency; x, y. and z are accelerations in each axis). The average 1-s value of the last 2 min of SVM_g values of each trial was used to establish cutpoints.

$$\mathrm{SVM} = \sum_{i=1}^{f} |\sqrt{x^2 + y^2 + x^2}|/(f) \qquad (1)$$

Metabolic measures. We measured the rates of oxygen consumption ($\dot{V}O_2$) and carbon dioxide production ($\dot{V}CO_2$) to determine the metabolic rate using a portable open circuit respirometry system (Oxycon Mobile, Yorba Linda, CA). Pediatric-specific masks were used on our subjects when necessary. The Oxycon Mobile provides valid measures of oxygen consumption across a range of exercise intensities (24) and has been used in calibration experiments with children (2,3,32). Before the experimental trials, we calibrated the system using gases with known concentrations. During each activity trial, expired gas data were averaged every 30 s. We allotted 6 min to ensure subjects reached steady state, which was defined as no significant increase in $\dot{V}O_2$ during the final 2 min and a RER <1.0. We then calculated the average $\dot{V}O_2$ and $\dot{V}CO_2$ (mL·s^{-1}) for the final 2 min of each trial.

Data and Statistical Analysis

To establish subject-specific resting metabolic rates, we used the Schofield equation for estimating resting energy expenditure (28). We then divided the measured $\dot{V}O_2$ value for each activity by the Schofield predicted resting value to determine MET values for each activity. Receiver operator characteristics (ROC) curves were generated using a leave-one-out (LOO) cross-validation to determine appropriate SVM_g values for cutpoints associated with sedentary (≤1.5 METs), light (>1.5–2.99 METs), moderate (3–5.99 METs), and vigorous (≥6 METs) activities. To generate these ROC curves, the last 2 min of SVM_g values (n = 120 values) for each activity for all subjects (less the left-out subject) were associated with the average MET value over the same time period of each trial. For each ROC curve, MET values were coded as a 0 or 1 according to the cutpoint being established (per SPSS methodology). For example, when the vigorous cutpoint was being established, a "1" was assigned to all vigorous activities, while a "0" was assigned to all activities not of vigorous intensity. The area under the ROC curve was calculated for sedentary, moderate, and vigorous activities, and cutpoint values were selected where the sum of the sensitivity and specificity was greatest. The sedentary and moderate cutpoints established the boundaries for light activity. The final set of cutpoints was established by averaging the values generated from each iteration. To examine the accuracy of the average cutpoints in estimating activity intensity, each left-out child then served as the test subject. This LOO process was repeated for all children. An average confusion matrix was constructed to examine how well the final set of cutpoints accurately classified activity on the left-out subject (i.e., the test subject). SPSS was used (Version 20; IBM, Somers, NY) for all statistical analysis.

Application to a Free-Living, Independent Sample

To determine estimates of daily PA using the cutpoints as well as to examine the effects of various bout duration minimums, we applied the cutpoints to an independent sample of free-living data from the Intervention of PhysicaL Activity in Youth (IPLAY) Study. IPLAY is a multischool intervention that aims, in part, to examine the effects of playground renovations and recess curriculum on levels of PA in elementary school students. The subsample to which the cutpoints were applied comprised 47 elementary school children (one first-, third-, and fifth-grade class). Table 1 includes the descriptive statistics for the IPLAY data sample. GENEActiv ACCs were attached to each child's nondominant

TABLE 2. Descriptive statistics for each laboratory calibration activity.

Activity	n	Mean SVM_g	Mean $\dot{V}O_2$ ($mL \cdot kg^{-1} \cdot min^{-1}$)	Mean METs
Resting	23	0.050 ± 0.05	4.89 ± 0.6	1.00 ± 0.0
Coloring	23	0.112 ± 0.04	7.68 ± 1.6	1.57 ± 0.2
Legos	23	0.234 ± 0.06	8.06 ± 2.4	1.65 ± 0.5
Wii tennis	23	0.353 ± 0.13	13.55 ± 4.2	2.78 ± 0.8
Wii boxing	22	1.290 ± 0.46	18.72 ± 5.4	3.90 ± 1.2
Slow walking, 0.75 $m \cdot s^{-1}$	23	0.296 ± 0.16	14.16 ± 2.1	2.91 ± 0.4
Pref walking, 1.25 $m \cdot s^{-1}$	22	0.381 ± 0.07	18.39 ± 2.6	3.76 ± 0.5
Jogging, 1.75 $m \cdot s^{-1}$	23	1.277 ± 0.24	30.33 ± 5.7	6.24 ± 1.2
Running, 2.25 $m \cdot s^{-1}$	22	1.594 ± 0.20	37.74 ± 2.8	7.83 ± 0.9

Values are reported as mean ± SD.
Pref, average preferred walking speed; SVM_g, gravity-subtracted signal vector magnitude.

wrist and secured using a plastic, nonelastic, hospital-type band (Wristbands MedTech USA, Orlando, FL). The devices were worn for 6 d (including two weekend days), and data were recorded at a sampling frequency of 75 Hz. As in our calibration experiment, a custom MATLAB program was used to process the data into 1-s average SVM_g (see Methods, Equation 1), to analyze the ACC data and to define custom intervals for the time periods of interest (e.g., weekday, school day, recess). We examined a single weekday (midweek, to avoid atypical activities including field day and field trips), with an outcome measure of minutes of MVPA, based on bout lengths of 1, 2, 3, 5, 10, 15, and 60 s using cutpoints established in the laboratory-based protocol. In an attempt to better understand how much of children's activity is vigorous in nature, we then examined vigorous PA (VPA) based on bout lengths of 1, 2, 3, and 5 s. The bout analysis was done using a custom MATLAB program whereby consecutive seconds of data above the moderate (or vigorous) threshold were summed. Independent-sample t-tests were conducted to determine differences in age, height, and weight between the calibration sample and the IPLAY sample population. ANOVA with Tukey *post hoc* was used to examine differences in the minutes of MVPA when applying the different bout durations to the IPLAY sample. SigmaPlot (Version 11.0, San Jose, CA) was used for the statistics involving the IPLAY subsample.

RESULTS

ACC data and oxygen consumption. Descriptive statistics for the ACC SVM_g output, $\dot{V}O_2$ value, and MET value for each activity trial are listed in Table 2. The average cutpoints resulting from the LOO validation, as well as the average values for area under the curve (AUC), sensitivity, and specificity, are listed in Table 3. Sedentary, moderate, and vigorous cutpoint values were 0.190, 0.314, and 0.998, respectively. We encountered an error with one subject's ACC data, and one subject was unable to complete two trials; therefore, the subject sample size ranged from 22 to 23. When examining the ability of the cutpoints to accurately classify intensity, we found that sedentary and vigorous activities were classified with relatively good accuracy (83.3% and 88.7%, respectively), whereas light and moderate activities were less accurately classified (27.6% and 41%, respectively; Table 4). When grouping sedentary and light activities together (SL), as well as moderate and vigorous activities (MV), classification accuracies improved (SL = 75.2%, MV = 69.7%). AUC, sensitivity, and specificity are listed in Table 3.

Application to IPLAY subsample. Using a 1-s bout, mean daily MVPA in the free-living sample was estimated to be 308.2 min. Results of the bout duration analysis when applied to the free-living sample are displayed in Figures 1 (MVPA) and 2 (VPA). As hypothesized, total accumulated minutes of MVPA and VPA decreased as the length of the bout increased. When a 60-s minimum bout duration was used, the average MVPA was 14.3 min. One-second MVPA values were significantly greater than 5-, 10-, 15-, and 60-s bouts but were not significantly different from 2- and 3-s bouts. VPA decreased from 32.7 min when assessing activity using a 1-s bout minimum to 12.4 min when using a 5-s minimum bout duration. One-second VPA values were significantly different than 3- and 5-s bouts. When bout length was increased from 1 to 5 s, MVPA decreased by ~32% while VPA decreased by ~60%.

DISCUSSION

The primary aim of this study was to establish cutpoints for sedentary, light, moderate, and vigorous activities in 6- to 11-yr-old children using a wrist-mounted GENEActiv ACC sampling at 75 Hz. SVM_g cutpoints were 0.190g, 0.314g, and 0.998g for sedentary, moderate, and vigorous activities, respectively. The cutpoints distinguished MVPA with reasonable accuracy (~70%), supporting our hypothesis.

TABLE 3. ROC-established cutpoints (SVM_g) and corresponding sensitivity and 1 − specificity values per 1-s epoch (n = 23).

	1-s Epoch Cutpoint, Mean ± SD	Sensitivity	Specificity	Area Under ROC Curve	SE	P
Sedentary	0.190 (0.0016)	0.970	0.876	0.956	0.015	<0.001
Light	NA	NA	NA	NA	NA	NA
Moderate	0.314 (0.0004)	0.910	0.873	0.946	0.015	<0.001
Vigorous	0.998 (0.0024)	0.949	0.853	0.940	0.016	<0.001

NA, not available; ROC, receiver operator characteristics; SVM_g, gravity-subtracted signal vector magnitude.

APPLIED SCIENCES

TABLE 4. Average confusion matrix (n = 23) indicating the ability of cutpoints to accurately classify activities (% accurately classified).

	Sedentary	Light	Moderate	Vigorous
Sedentary	83.3	5.8	4.7	0
Light	47.6	27.6	21.1	3.6
Moderate	5.5	24.9	41	28.7
Vigorous	0.7	1.2	9.3	88.7

Columns indicate actual activity, while rows indicate predicted activity.

Using these cutpoints, accumulated minutes of daily MVPA were estimated to be ~300 min in an independent free-living sample and, as hypothesized, decreased with increasing bout duration.

Few published studies have attempted to establish wrist-mounted cutpoints using the GENEActiv ACC (10,21). A recent study by Phillips et al. (21) established activity intensity cutpoints for the wrist-mounted GENEActiv in children using a methodology similar to our study. Age group, activity choice, and much of the postprocessing methodology were similar. By multiplying our cutpoint values by the sampling frequency, we were able to compare our values to those established by Phillips et al. Values established in this study (sedentary = 14.25*g*, moderate = 23.4*g*, vigorous = 74.85*g*) are greater than those reported by Phillips et al. (sedentary = 7*g*, moderate = 20*g*, vigorous = 60*g*). However, the AUC values from the ROC curves were similar, suggesting similar classification performance (21). The most notable difference between our values and those established by Phillips et al. is in the sedentary cutpoint (14.25*g* vs 7*g*). This likely reflects our selection of sedentary tasks involving the use of the hands/wrist (i.e., coloring and legos), whereas the sedentary activities performed by Phillips et al. included lying, sitting, and DVD watching (minimal wrist movement). The similarity of the moderate cutpoints is encouraging and not surprising given the similarity of the moderate-intensity activities (e.g., walking). If the cutpoints of Phillips et al. were applied to our free-living, independent

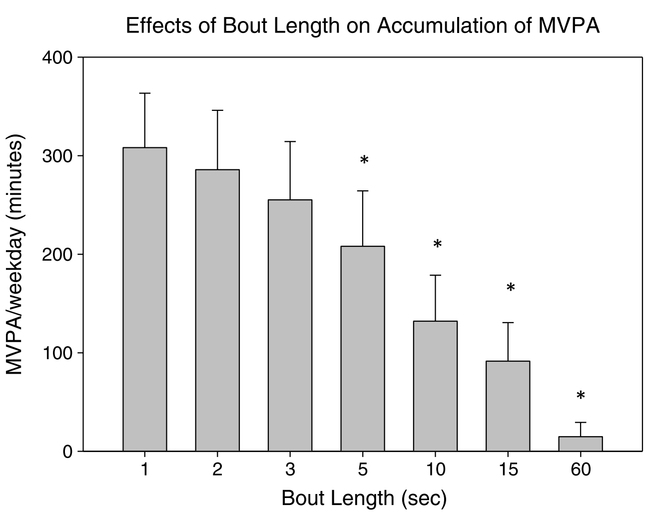

FIGURE 1—Effects of bout length duration on the accumulation of minutes of moderate–vigorous physical activity during a single weekday (mean, SE). *Significantly different compared to 1-s bout (P < 0.05).

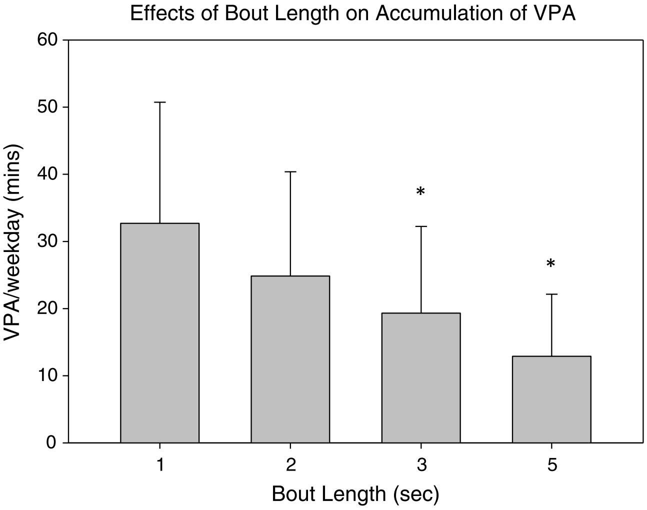

FIGURE 2—Effects of bout length duration on the accumulation of minutes of vigorous physical activity during a single weekday (mean, SE). *Significantly different compared to 1-s bout (P < 0.05).

sample, fewer minutes of sedentary activity would be classified, but a similar or slightly greater number of minutes of MVPA would be observed compared to what we report here. As no study has yet applied the cutpoints of Phillips et al. to an independent free-living sample, we are not able to compare the ability of these cutpoints to quantify accumulated minutes of MVPA.

We examined the ability of the cutpoints to accurately classify activity in our participants by creating a confusion matrix based on the calibration data. To our knowledge, no groups have used a similar methodology to quantify how well cutpoints are able to distinguish activity intensity levels. Our results demonstrate that sedentary and vigorous activities are classified with good accuracy (83.3% and 88.7%, respectively). Although the classification accuracies of light and moderate activities are not as accurate, by grouping SL and MV activities, accuracies improve substantially (SL = 75.2%, MV = 69.7%). While these values indicate that up to 30% of activity may be misclassified, a similar percentage of activities are likely classified too low as those classified too high. Importantly, these classification accuracies likely represent a best case given we used the same subjects for cutpoint determination and classification testing. The challenge associated with accurately classifying activity intensity using ACCs is similar to that of using these devices to classify free-living activities and predict metabolic rate. Laboratory-based validations of activity classification report good accuracies (>90% for general classes of activities) (36), a significant improvement compared to those reported here. The better activity versus intensity classification accuracies are likely due to the more sophisticated classification methodology (e.g., machine learning with multiple features) used in activity classification and suggests that intensity classification could improve with such approaches. Calibration studies that have used linear regression to estimate metabolic rate or energy expenditure report R^2 values ranging from 0.35 to 0.84 (13,20,22), indicating that a substantial portion of the variance in the relationship (16%–65%) is not explained by ACC output. Although the ability of ACCs to accurately estimate activity intensity may improve as additional calibration studies are conducted, using acceleration data to classify activities and estimate activity-specific energy expenditure potentially offers a promising alternative use of this data. However, even robust classifiers are only as good as the data used to develop them, and our results suggest the possibility that the way children move in a calibration experiment is not the way they move in a free-living study. If true, classifiers intended to quantify intensity and activity in children may improve when free-living data are used to develop them.

When we applied the cutpoints to an age-, height-, and weight-matched free-living sample of children wearing GENEActiv ACCs collecting data at 75 Hz and processed identically to the calibration study, estimates of minutes of MVPA (mean MVPA = 308 min) were much larger than those typically reported for children of this age. However, most published studies reporting objectively measured average daily minutes of MVPA in children have been conducted using 1-min epochs (31). No studies have quantified MVPA in children using a 1-s epoch, although many have acknowledged the need to do so (5,9,17,25). Studies that have used hip-mounted devices recording 2-s epochs report MVPA ranging from 80 (5) to 160 $\text{min}\cdot\text{d}^{-1}$ (27). Possible explanations for the wide range of accumulated MVPA include significant variation in children's activity levels, seasonal variations, differences in wear time, device location, and on-board processing of the data (e.g., data filtering). Interestingly, in a series of studies conducted by Sleap and Warburton that involved direct observation of 5- to 11-yr-old children using the Children's Physical Activity Form (which measures children's PA every 3 s) (19), an average of 122 min of MVPA was accumulated per observation period (~418 min), more than double the suggested daily guideline (29). Approximately half of the observed time during school playtimes and one quarter of the time outside of school was observed to be MVPA. If we extrapolate these percentages to a full day, assuming an estimated 8 h of out-of-school time (6–8 a.m. and 3–9 p.m.) and 2 h of school playtime, including break time, lunch, recess, PE, before, and after school, estimates of nearly 200 min of MVPA may be observed (25% × 8 h out-of-school time each day + 50% × 2 h of school playtime = 3 h of MVPA). These direct observation studies support our finding that children engage in substantial amounts of daily MVPA. However, the modest intensity classification accuracy reported here, combined with the MVPA estimates that exceed other published estimates, suggests that wrist-mounted ACC cutpoint-derived estimates of MVPA may, at present, be better suited for measuring changes in MVPA rather than as an accurate measure of MVPA.

There have been few attempts to explore how much continuous MVPA children perform. Studies examining the average bout duration of activity in various intensity levels suggest that children perform MVPA for short periods of time. Using a 2-s epoch, a study conducted by Baquet et al. (5) found the average duration of moderate (3–6 METs), vigorous (6–9 METs), and very high (>9 METs) bouts of PA to be 9, 4.7, and 3.9 s, respectively. Additional studies exploring children's activity patterns at 2-s epochs suggest similar trends (26,30). Our results are consistent with these findings. When increasing minimum bout duration from 1 to 5 s, estimates of vigorous activity decreased by over 60% (32.7 vs 12.9 min of VPA). This information is useful in the design of interventions aiming to increase children's MVPA. A novel strategy may be to focus on increasing the duration of short bouts (e.g., extending PA from one to five consecutive seconds) rather than encouraging prolonged PA (i.e., anything longer than ~10–15 consecutive seconds). An alternative strategy may be to increase the number of (rather than the duration of) short bouts. These may be more effective strategies for accumulating PA, given that it may mimic children's natural PA patterns. Future studies that investigate the dose–response of MVPA bout duration on health outcomes are also needed to establish effective PA guidelines for children (31).

We acknowledge that our findings indicating that children accumulate >300 $\text{min}\cdot\text{d}^{-1}$ of MVPA are much greater than estimates in many published studies. One potential explanation was our use of 3 and 6 METs as the cutoffs for moderate and vigorous activities, respectively. Some groups have suggested that 4 and 7 METs are more appropriate cutoffs for children (31). However, this recommendation seems to have been based on using a standard resting metabolic rate value of 3.5 $\text{mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$, rather than using age-specific resting metabolic rate estimates. To examine the difference in MVPA accumulation using 4 versus 3 METs, we created a 4-MET cutpoint and applied it to our data. The moderate cutpoint changed by approximately 22% (from 0.314 to 0.400), and estimates of MVPA changed by 27% (from 308 to 225 min). Another contributing factor to the amount of MVPA reported here could be that light PA is being misclassified as MVPA. While we acknowledge this possibility, it is also likely that some MVPA is also being misclassified at light. An additional explanation for our free-living estimates may be that the activities in the calibration study were not performed the same way in a free-living setting, making it difficult for the correct intensity to be estimated. Of course, it is also possible that our calibration activities are not representative of typical children's activities. This points to the critical need for a taxonomy of child-specific activities from which to select to more appropriately calibrate devices. Clearly, future studies are needed that validate the amount of MVPA children accumulate during a day using direct observation or other technique with an equivalent short sampling interval.

The move toward the collection of raw acceleration data highlights the need to standardize methodologies for data processing, establishing cutpoints, classifying PA, and quantifying duration and intensity of PA (35). Potential ways by which to do so include standardizing 1) the activities/speeds selected for validation studies; 2) the method selected for processing the data, including the frequency with which data are collected and the filtering applied to the raw signal; 3) the analysis method for deriving and validating the cutpoints or classifier (e.g., ROC vs machine learning); and 4) the ACC and device location selected (e.g., wrist vs hip) (14). To facilitate comparisons across studies, in the Supplemental Digital Content (SDC), we have included sample .docx versions of the MATLAB code used to process the raw .bin file (SDC 1, http://links.lww.com/MSS/A371 binread.docx—reads the bin file; SDC 2, http://links.lww.com/MSS/A372 convertbin.docx—converts the file from a .bin to a .csv file; and SDC 3, http://links.lww.com/MSS/A373 filterbin.docx—applies the band pass filter). We have also included a sample output file from the calibration study (SDC 4, http://links.lww.com/MSS/A379 Sample1_Filtered_Timestamp.xlsx) along with the time codes used in the calibration trial (SDC 5, http://links.lww.com/MSS/A374 Sample1_ActivityTimes.xlsx). Samples of the raw .bin data files may be obtained by contacting the corresponding author.

Limitations. Our study is not without limitations. First, our sample size for the calibration trial was relatively small (n = 24) and not widely ethnically diverse, which may limit the generalizability of our findings. In addition, we elected to use a prediction equation that is validated for children (28) rather than using our measured values for resting EE. Because we did not measure resting EE under the stringent conditions required for a true resting measurement, we believe using the well-established Schofield equation would minimize any potential error associated with the baseline resting value. The degree to which the established cutpoints can be applied to a population is dependent on the similarity in age, size, behavioral patterns, and activities undertaken between the two populations (35). The subsample of children to which our cutpoints were applied was not significantly different in terms of age, sex, height, or weight from the calibration sample. However, there may be behavioral differences between the two populations that we are unable to detect. Understanding the degree to which behavioral differences affect the estimated levels of PA is important to accurately quantify PA in children. In addition, the differences between the activities selected for the calibration study and the actual activities undertaken during free-living activity are important considerations and should be addressed in future calibration studies. Finally, although we conducted a LOO cross-validation to create the confusion matrix (Table 4), we did not apply the cutpoints to an independent sample. As a whole, the field of PA monitoring acknowledges that this is an important future step in assessing the accuracy of both intensity and activity classifiers (8).

Conclusions. Using ROC curves, the calibration of the wrist-mounted GENEActiv in elementary school-age children resulted in intensity cutpoints of 0.019, 0.314, and 0.998 SVM_g for sedentary, moderate, and vigorous activities, respectively. MVPA intensity classification accuracy was moderately good (~70%). When applied to a free-living data set, we estimated 308 $min \cdot d^{-1}$ of MVPA, suggesting that children move frequently and intermittently throughout the day. As we move toward raw data collection, researchers will need to explore how to interpret the physiological meaningfulness of these very short bouts of activity.

The authors would like to thank Valerie Ward, Paige Kyle, Emily Thompson, and other members of the Physical Activity Energetics and Mechanics Lab at Colorado State University, as well as the IPLAY study team, for their help in collecting these data. The IPLAY study is funded by the National Institutes of Health, NICHD/NCI/NIDDK R01HD057229.

The authors declare no conflict of interest. Results of the present study do not constitute endorsement by the American College of Sports Medicine.

REFERENCES

1. Antonsson EK, Mann RW. The frequency content of gait. *J Biomechanics*. 1985;18(1):39–47.
2. Arvidsson D, Slinde F, Larsson S, Hulthen L. Energy cost of physical activities in children: validation of sensewear armband. *Med Sci Sports Exerc*. 2007;39(11):2076–84.
3. Arvidsson D, Slinde F, Larsson S, Hulthen L. Energy cost in children assessed by multisensor activity monitors. *Med Sci Sports Exerc*. 2009;41(3):603–11.
4. Bailey RC, Olson J, Pepper SL, Porszasz J, Barstow TJ, Cooper DM. The level and tempo of children's physical activities—an observational study. *Med Sci Sports Exerc*. 1995;27(7):1033–41.
5. Baquet G, Stratton G, Van Praagh E, Berthoin S. Improving physical activity assessment in prepubertal children with high-frequency accelerometry monitoring: a methodological issue. *Am J Prev Med*. 2006;44(2):143–7.
6. Berman N, Bailey R, Barstow TJ, Cooper DM. Spectral and bout detection analysis of physical activity patterns in healthy, prepubertal boys and girls. *Am J Hum Biol*. 1998;10(3):289–97.
7. Butte NF, Ekelund U, Westerterp KR. Assessing physical activity using wearable monitors: measures of physical activity. *Med Sci Sports Exerc*. 2012;44(1 Suppl 1):S5–12.
8. Crouter SE, Horton M, Bassett DR Jr. Validity of ActiGraph child-specific equations during various physical activities. *Med Sci Sports Exerc*. 2013;45(7):1403–9.
9. Edwardson CL, Gorely T. Epoch length and its effect on physical activity intensity. *Med Sci Sports Exerc*. 2010;42(5):928–34.
10. Esliger DW, Rowlands AV, Hurst TL, Catt M, Murray P, Eston RG. Validation of the GENEA accelerometer. *Med Sci Sports Exerc*. 2011;43(6):1085–93.
11. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. *J Sports Sci*. 2008;26(14):1557–65.
12. Freedson P, Bowles HR, Troiano R, Haskell W. Assessment of physical activity using wearable monitors: recommendations for monitor calibration and use in the field. *Med Sci Sports Exerc*. 2012;44:S1–4.
13. Freedson P, Pober D, Janz KF. Calibration of accelerometer output for children. *Med Sci Sports Exerc*. 2005;37(11 Suppl):S523–30.
14. Intille SS, Lester J, Sallis JF, Duncan G. New horizons in sensor development. *Med Sci Sports Exerc*. 2012;44:S24–31.
15. Jago R, Zakeri I, Baranowski T, Watson K. Decision boundaries and receiver operating characteristic curves: new methods for

determining accelerometer cutpoints. *J Sports Sci*. 2007;25(8): 937–44.
16. Kram R, Griffin TM, Donelan JM, Chang YH. Force treadmill for measuring vertical and horizontal ground reaction forces. *J Appl Physiol*. 1998;85(2):764–9.
17. McClain JJ, Abraham TL, Brusseau TA Jr, Tudor-Locke C. Epoch length and accelerometer outputs in children: comparison to direct observation. *Med Sci Sports Exerc*. 2008;40(12): 2080–7.
18. Nilsson A, Ekelund U, Yngve A, Sjostrom M. Assessing physical activity among children with accelerometers using different time sampling intervals and placements. *Pediatr Exerc Sci*. 2002;14: 87–96.
19. O'Hara NM, Baranowski T, Simons-Morton BG, Wilson BS, Parcel G. Validity of the observation of children's physical activity. *Res Q Exerc Sport*. 1989;60(1):42–7.
20. Pfeiffer KA, McIver KL, Dowda M, Almeida M, Pate RR. Validation and calibration of the Actical accelerometer in preschool children. *J Med Sci Sports Exerc*. 2006;38(1):152–7.
21. Phillips LR, Parfitt G, Rowlands AV. Calibration of the GENEA accelerometer for assessment of physical activity intensity in children. *J Sci Med Sport*. 2013;16(2):124–8.
22. Puyau MR, Adolph AL, Vohra FA, Butte NF. Validation and calibration of physical activity monitors in children. *Obesity Research*. 2002;10(3):150–7.
23. Rodgers AB. 2008 Physical Activity Guidelines for Americans. Department of Health and Human Services 2008. Available from: Department of Health and Human Services.
24. Rosdahl H, Gullstrand L, Salier-Eriksson J, Johansson P, Schantz P. Evaluation of the Oxycon Mobile metabolic system against the Douglas bag method. *Eur J Appl Physiol*. 2010; 109(2):159–71.
25. Rowlands AV. Accelerometer assessment of physical activity in children: an update. *Pediatr Exerc Sci*. 2007;19(3):252–66.
26. Rowlands AV, Pilgrim EL, Eston RG. Patterns of habitual activity across weekdays and weekend days in 9–11-year-old children. *Prev Med*. 2008;46(4):317–24.
27. Rowlands AV, Pilgrim EL, Eston RG. Seasonal changes in children's physical activity: an examination of group changes, intra-individual variability and consistency in activity pattern across season. *Ann Hum Biol*. 2009;36(4):363–78.
28. Schofield WN. Predicting basal metabolic rate, new standards and review of previous work. *Hum Nutr Clin Nutr*. 1985;39(Suppl 1):5–41.
29. Sleap M, Warburton P. Physical activity levels of 5–11-year-old children in England: cumulative evidence from three direct observation studies. *Int J Sports Med*. 1996;17(4):248–53.
30. Stone MR, Rowlands AV, Eston RG. Characteristics of the activity pattern in normal weight and overweight boys. *Prev Med*. 2009; 49(2–3):205–8.
31. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc*. 2008;40(1):181–8.
32. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of accelerometer cut points for predicting activity intensity in youth. *Med Sci Sports Exerc*. 2011;43(7):1360–8.
33. Trost SG, Wong WK, Pfeiffer KA, Zheng Y. Artificial neural networks to predict activity type and energy expenditure in youth. *Med Sci Sports Exerc*. 2012;44(9):1801–9.
34. Ward V, Schaefer C, Lampe S, et al. Achieving excellent compliance in multi-day free-living activity monitoring in elementary school students: the IPLAY study. *Med Sci Sports Exerc*. 2011; 43(5):S486.
35. Welk GJ. Principles of design and analyses for the calibration of accelerometry-based activity monitors. *Med Sci Sports Exerc*. 2005; 37(Suppl 11):S501–11.
36. Zhang SY, Rowlands AV, Murray P, Hurst TL. Physical activity classification using the GENEA wrist-worn accelerometer. *Med Sci Sports Exerc*. 2012;44(4):742–8.