Validation of Wearable Monitors for Assessing Sedentary Behavior

SARAH KOZEY-KEADLE1, AMANDA LIBERTINE1, KATE LYDEN1, JOHN STAUDENMAYER2, and PATTY S. FREEDSON1

1Department of Kinesiology, University of Massachusetts, Amherst, MA; and 2Department of Math and Statistics, University of Massachusetts, Amherst, MA

ABSTRACT
KOZEY-KEADLE, S., A. LIBERTINE, K. LYDEN, J. STAUDENMAYER, and P. S. FREEDSON. Validation of Wearable Monitors for Assessing Sedentary Behavior. Med. Sci. Sports Exerc., Vol. 43, No. 8, pp. 1561–1567, 2011. Purpose: A primary barrier to elucidating the association between sedentary behavior (SB) and health outcomes is the lack of valid monitors to assess SB in a free-living environment. The purpose of this study was to examine the validity of commercially available monitors to assess SB. Methods: Twenty overweight (mean ± SD: body mass index = 33.7 ± 5.7 kg m⁻²) inactive, office workers age 46.5 ± 10.7 yr were directly observed for two 6-h periods while wearing an activPAL (AP) and an ActiGraph GT3X (AG). During the second observation, participants were instructed to reduce sitting time. We assessed the validity of the commonly used cut point of 100 counts per minute (AG100) and several additional AG cut points for defining SB. We used direct observation (DO) using focal sampling with duration coding to record either sedentary (sitting/lying) or nonsedentary behavior. The accuracy and precision of the monitors and the sensitivity of the monitors to detect reductions in sitting time were assessed using mixed-model repeated-measures analyses. Results: On average, the AP and the AG100 underestimated sitting time by 2.8% and 4.9%, respectively. The correlation between the AP and DO was R² = 0.94, and the AG100 and DO sedentary minutes was R² = 0.39. Only the AP was able to detect reductions in sitting time. The AG 150-counts-per-minute threshold demonstrated the lowest bias (1.8%) of the AG cut points. Conclusions: The AP was more precise and more sensitive to reductions in sitting time than the AG, and thus, studies designed to assess SB should consider using the AP. When the AG monitor is used, 150 counts per minute may be the most appropriate cut point to define SB. Key Words: MEASUREMENT, DIRECT OBSERVATION, ACCELEROMETER, SITTING TIME

Sedentary behavior, defined as energy expenditure between 1 and 1.5 METs while sitting or lying, is detrimental to one’s health (15). Independent of physical activity status, there are positive associations between sedentary behavior and risk of obesity (21), metabolic syndrome (3), type 2 diabetes (9), and mortality (1,11). Despite these observations, we lack validated instruments to measure sedentary behavior (16). The majority of sedentary behavior research uses self-report questionnaires including surrogate measures such as time spent watching TV (2). However, no self-report measure comprehensively assesses all components of sedentary behavior. In particular, patterns of inactivity such as breaks <5 min or changes in sedentary behavior are challenging to measure with a self-report instrument.

In response to the limitations of self-report instruments, researchers have begun to use objective measures including pedometers and accelerometers to quantify sedentary behavior. Five thousand steps per day defines the upper boundary for sedentary behavior using a pedometer, but this definition does not distinguish between sitting and standing time, nor does it describe patterns of inactivity within a day (22). As a result, researchers primarily use accelerometer-based activity monitors to assess sedentary behavior. In studies that use the ActiGraph (AG; ActiGraph LLC, Pensacola, FL) activity monitor, a sedentary minute is defined as one when the monitor output is <100 counts per minute (14). Such studies have shown a robust relationship between objectively measured sedentary behavior and health outcomes (7,8). Although widely used, the 100-counts-per-minute cut point (AG100) was not empirically derived. In addition, the AG monitor is a single hip-mounted device that may not be able to distinguish postures (e.g., sitting vs standing). For example, the AG monitor output for standing activities, such as folding laundry and washing dishes, can be near or below 100 counts per minute (12), and these activities are not sedentary. In general, the ability of this monitor to distinguish between sedentary time and light-intensity activity time is not known. The activPAL (AP; Physical Activity Technologies, Glasgow, Scotland) is a promising tool designed specifically to measure free-living activity. It has the ability to differentiate among postures and classify an individual’s activity into time sitting, standing, and stepping. This device has been validated...
in the laboratory compared with a criterion measure (direct observation (DO)) and was recently found to be 100% accurate for measuring sitting, standing, and walking (5,19). However, the AP has not been validated in a free-living setting compared with DO. A recent study examined the convergent validity of the AG and the AP and reported that, on average, the AG recorded 132 min more sedentary time than the AP for 15 h (6). In this study, a criterion measure was not used, and thus, it cannot be determined which monitor was more accurate.

These activity monitors have not been validated for assessment of sedentary behavior in a free-living environment compared with a criterion measure. Therefore, the primary aim of this study was to validate the AG100 and the AP monitor for assessing sedentary behavior. We validated the monitors in two ways: 1) assessing the difference between monitor estimates and DO measures of sedentary behavior and 2) examining monitor performance in detecting reductions in sedentary behavior among inactive individuals. A secondary aim was to determine whether the AG100 is the most appropriate cut point for the AG. We compared the validity of the AG100 to other count cut points ranging from 50 counts per minute (AG50) to 250 counts per minute (AG250) using DO as the criterion method.

METHODS

Eligibility and recruitment. Participants were recruited from the University of Massachusetts, Amherst, and local communities via fliers and word of mouth. Eligible participants were at least 25 yr, overweight or obese (body mass index (BMI) >25 kg m$^{-2}$), and inactive, which was defined as participating in <3 d wk$^{-1}$ of moderate physical activity for 20 min per session during the preceding 6 months. Participants were employed in an occupation where most of the work day was spent sitting.

Eligible participants reported to the University of Massachusetts and signed an informed consent document that was approved by the University of Massachusetts Institutional Review Board. Participants then completed a physical activity readiness questionnaire, a health history questionnaire, and a physical activity status questionnaire. After the consenting process, height and weight (to the nearest 0.1 kg) were measured using a floor scale/stadiometer (Detecto, Webb City, MO) while participants wore a thin layer of clothing and no shoes. The sample included five males and 15 females. Mean ± SD age of participants was 46.5 ± 10.7 yr. Mean ± SD body mass index was 33.7 ± 5.7 kg m$^{-2}$.

Procedures. Participants completed two 7-d conditions. The first condition was a baseline measurement where participants were instructed to wear the activity monitors during all waking hours each day. During both conditions, participants were directly observed in their free-living environment for one 6-h period.

Strategies to reduce sitting. At the end of the 7-d sedentary condition, participants were given recommendations to increase their time standing and decrease their time sitting. They were provided with detailed information about the health risks associated with sedentary behavior and the benefits associated with increasing light-intensity activity. They were given examples and strategies for decreasing sedentary time and accumulating light-intensity activity (e.g., standing during all commercials while watching television, taking a 5-min “standing/walking” break each hour at work). To help facilitate compliance, participants were given daily and hourly checklists of tasks to complete. The checklists helped participants self-monitor their compliance and also served as regular reminders for participants to break up their sitting time. During the active condition, participants were given a pedometer step goal of at least 7500 steps per day. This step goal has been designated as the lower boundary of “somewhat active” behavior (22).

Criterion measure: DO. Participants were observed for six consecutive hours, once per condition. The majority of observations took place during participants’ working hours. A custom DO program was developed for a personal digital assistant (PDA; Palm Tungsten E2; Palm, Inc., Sunnyvale, CA). The PDA was synchronized with the activity monitors before each data collection session. Three researchers completed DO training that included review of a training manual, 2 h of training videos, and DO practice sessions with the PDA for a minimum of 12 h. After the training, subjects completed a testing video that was 25 min in duration and included 20 different video clips, each containing various postures and activities. Before data collection, researchers were required to correctly classify 90% of the body positions, intensity levels, and duration of activities throughout the training video.

Focal sampling and duration coding were used, with trained data collectors coding the real-time occurrence of the five activity categories, body positions, and intensities described below:

1. Lying: Individuals were flat on their backs (horizontal).
2. Sitting: Individuals had some of their body weight supported by the buttocks or thighs. The upper body was not parallel to the ground. If they were kneeling, they were coded based on the thigh position (i.e., if the thigh is parallel to the ground, sitting was selected).
3. Standing still: Individuals were standing with little or no contribution from the upper body. They were not carrying a load >1 kg. Standing still included talking with hand gestures, looking at something, or waiting in a line.
4. Standing still with upper body movement: Individuals were upright with some contribution from the upper body that causes an increase in energy expenditure...
(holding a load >1 kg, filing papers, or doing a task that requires the arms). The purpose of the activity had to include the upper body.

5. Standing/moving: Individuals were engaging in activities that are of light intensity (<3 METs; e.g., walking at a speed <2.5 mph and not be carrying a load). These activities included movements around an office or a home but not for locomotion (e.g., traveling between one place and another).

6. Moving moderate: Individuals were engaging in activities >3 METs. Examples include walking >2.5 mph, gardening, vacuuming, and carrying a load.

7. Moving vigorous: Individuals were engaging in activities >6 METs. This typically involves purposeful exercise including jogging, walking briskly uphill, and sporting activities.

Total sedentary time was determined by summing/totaling the amount of time spent in lying and sitting body positions from the DO coding system. Any other body positions or postures were not considered sedentary behaviors.

**Activity monitors.** The AP is a small (2.0 × 1.4 × 0.3 inches) and light (20.1 g) single-unit accelerometer device worn on the mid-thigh on right leg (attached by non-allergenic adhesive tape) and uses accelerometer-derived information about thigh position to estimate time spent in different body positions (horizontal = lying or sitting; vertical = standing) in 15-s epochs. When the participant was stepping, the device measured step cadence and number of steps. The AP output of time spent sitting/lying was defined as sedentary behavior.

The AG (model GT3X) is a small (1.5 × 1.44 × 0.7 inches) and light (28 g) triaxial accelerometer that was secured to the right side of the hip using an elastic belt. Firmware version 2.1.0 was used, and the low-frequency extension was selected. The monitor was initialized to record vertical acceleration in 1-s epochs. Sedentary time was defined as the sum of the minutes with corresponding monitor data. On average, each participant was observed for 346 min (5.8 h) per observation.

**Data cleaning.** To be included in the analysis, a participant was required to have simultaneous AG, AP, and DO data. Two participants wore the AP monitor upside down (data not included). One participant used a chair at work that supported the participant’s lumbar spine and resulted in a vertical thigh position (perpendicular to the floor) while the participant was seated. As a result, the observer was unsure how this should be coded, and sitting time was recorded as standing by the AP, thus the data from this participant were not included in the analysis. Of the 20 enrolled participants, 16 had valid data for both DO sessions and 19 participants had valid data for at least one DO session. This resulted in a total of 12,132 observation minutes with corresponding monitor data. On average, each participant was observed for 346 min (5.8 h) per observation.

**Statistical analyses.** To determine the validity of the AG100 and AP monitors, we performed two analyses. First, we compared the monitor estimates to the DO measures of sedentary time, and second, we evaluated the ability of the monitors to detect reductions in sitting time. We used a repeated-measures linear mixed model to determine the ability of the AG100 and AP to estimate sedentary time in free-living subjects compared with DO. Both accuracy (i.e., bias: the extent that each monitor overestimated or underestimated sedentary time) and precision (i.e., variability or random error: how far the estimate of sedentary minutes randomly fluctuate above and below its average value for each person on each day) were evaluated. We measured bias in units of minutes (monitor sedentary minutes – DO sedentary minutes) and as a percentage [(monitor sedentary minutes/DO sedentary minutes) – 1 × 100]. In both cases, positive biases indicated overestimates of sedentary behavior and negative values indicated underestimates of sedentary behavior. The percentage bias is useful because, for instance, a 10% bias could be applied to an observation time of 10 h (a 1-h overestimate) or an observation time of 70 h (a 7-h overestimate). We used correlation and confidence intervals (CI) as measures of precision. Higher precision was indicated by higher correlations and smaller CI.

The second method for validating the monitors was to evaluate if the monitors could detect changes in sedentary behavior between a sedentary and an active condition. Using the DO data, a subset of participants (n = 11) were identified who reduced their sitting time in the active condition compared with the sedentary condition. A repeated-measures linear mixed model was used to compare the differences in mean sitting time between conditions, and separate models were fit for DO, AP, and AG. Likelihood ratio tests were used to determine whether the differences were significant. The likelihood ratio test examined if the addition of condition as an independent variable resulted in a significantly better fit. If it did not, then the variability in the measurements was too large to statistically discern the changes in sedentary time within subjects. All statistical analyses were performed using R-software packages (www.r-project.org) (17). Significance levels were set at P < 0.05.

The secondary aim of the study was to determine whether the AG100 was the most accurate and precise cut point to assess sedentary behavior. The AG100 cut point was compared with cut points of 50, 150, 200, and 250 counts per minute. The analyses described above were repeated for each count cut point.

**RESULTS**

The directly observed data for time spent sedentary was normally distributed over the days and subjects within each
condition. The mean ± SD percent of directly observed time sedentary during the sedentary condition was 78.1% ± 16.5%, which is equivalent to 269.5 ± 60.9 sedentary minutes. For the active condition, the mean ± SD percent of observed time spent sedentary was 69.5% ± 11.2%, which is equivalent to 242.9 ± 43.0 sedentary minutes.

On average, both the AP and the AG100 underestimated sedentary time compared with DO. Figure 1 shows the bias in minutes and as a percentage. The AP bias was −7.7 min and SE was 2.5 min (95% CI = −12.5 to −2.9 min). The AG100 bias was −16.9 min and SE was 8.5 min (95% CI = −33.6 to −0.3 min). Using percent bias, the AP underestimated sitting time by 2.8% (SE = 1.0%, 95% CI = −4.7% to 0.9%), whereas the AG100 underestimated sitting time by 4.9% (SE = 3.4%, 95% CI = −11.6% to 1.8%). The results of the secondary aim analysis illustrate that the AG cut point with the lowest bias was AG150 (bias = −0.9 min, SE = 7.7 min, 95% CI = −15.9 to 14.1 min) (Fig. 1). The AG150 also had the lowest percent bias of 1.8% (95% CI = −5.3% to 8.9%). The percent biases and bias in minutes for AG50, AG200, and AG250 were higher than the commonly used AG100 (range = −22% to 17.8%; −60 to 32 min) (Fig. 1). Figure 2 is a modified Bland–Altman plot to illustrate the relationship between the DO and the AP percent of time sedentary ($R^2 = 0.94$), the DO and the AG100 ($R^2 = 0.39$), and the DO and AG150 percent of time sedentary ($R^2 = 0.40$).

Of the 16 participants with valid data at both observation periods, 11 reduced their sedentary time during the active condition compared with the sedentary condition. The smallest change in sitting time among the responders was a 2% reduction in sitting time during the active condition compared with the sedentary condition. In this subset of participants, the average percent of time sedentary was significantly different between conditions based on DO ($P < 0.01$; Fig. 3). According to DO, sedentary time was 83.7% ± 11.2% of the sedentary condition and 68.5% ± 11.4% of the active condition. Sedentary time was significantly different between conditions ($P < 0.01$); according to the AP, it was 79.5% ± 13.8% of the sedentary condition and 66.5% ± 10.2% of the active condition. The AG100 estimate of sedentary time was not significantly different between conditions ($P = 0.2$); it was equal to 70.5% ± 17.8% of the sedentary condition and 66.9% ± 11.9% of the active condition. Although the AG150 had the lowest bias for the AG monitor, it was not sensitive to reductions in sitting time between conditions ($P = 0.3$), nor were any other AG count cut points (Fig. 3).

**DISCUSSION**

As evidence accumulates that sedentary behavior is associated with premature mortality and chronic disease, it is imperative we have accurate measures of the time spent in sedentary behaviors (1,15). The major finding of the current study was that the AP is an accurate and precise monitor for measuring sedentary behavior and is sensitive to reductions in sitting time. Our results support the use of the AP in studies designed to determine the effects of sedentary behavior and changes in sedentary time on health outcomes. Another important finding was that the AG count cut point of 150 counts per minute was the most accurate AG cut point to define sedentary behavior. Using the previously defined sedentary cut point of 100 counts per minute for the AG monitor resulted in a significant underestimation of sitting time in our sample.

In this study, we report the bias and precision validation for each estimate of sitting time. Bias is the average difference...
between the estimate (monitor prediction) and the criterion (DO). The bias is commonly reported as it reflects the accuracy of the monitor and whether the monitor overestimates or underestimates sitting time. The AP had a slightly smaller bias (−2.8%) than the AG100’s bias (−4.9%), but these were not statistically different. Although bias is an important measure, when differences in sitting time before and after intervention are considered, the biases cancel each other. Thus, bias does not affect the sensitivity of the monitor to detect changes after an intervention. In contrast, precision (i.e., variability or random error) is of vital importance in the application to intervention trials. The higher precision of the AP compared with

FIGURE 2—Modified Bland–Altman plots of the relationship between DO and AP and AG estimates of percent time sedentary. The least squares regression line is dotted and the line at zero is dashed.

FIGURE 3—Sensitivity of monitors in distinguishing between sedentary and active conditions. *Difference between conditions is significant at $P < 0.05$. 
the AG (smaller SE, higher correlation) results in higher statistical power, more reliability, and smaller sample size requirements. This was illustrated in this study when we examined the sensitivity of the monitors to detect changes between conditions where only the AP could detect the reductions in sedentary behavior (Fig. 3).

A large bias and low precision also impairs the ability to identify a dose–response relationship between a sitting time and health outcomes. Data from the National Health and Nutrition Examination Survey (NHANES) using the AG100 to define sedentary behavior reported that adults spend 55% of their waking hours in sedentary behavior (14). Our results suggest sitting time was underestimated in the NHANES sample by approximately 4.9%, equivalent to 35 min during a 14-h day (14). Although this is a potentially important underestimation, it is a systematic error that could be corrected using a measurement error model. The wide confidence intervals of the AG100 are a more critical issue in our study because it reflects large random error, which cannot be corrected with a model. If we apply the estimates from our confidence intervals to the waking day (14 h) in the NHANES sample, the random error is between a 97-min underestimation and a 15-min overestimation of sedentary time. This nearly 2 h of random error is, by definition, unpredictable and leads to challenges in identifying doses of sedentary behavior that are detrimental to one’s health. Although the low precision of the AG monitor in measuring sedentary behavior is concerning, studies using this monitor have reported positive associations with sedentary behavior and disease risk (8). Therefore, future studies using a more accurate and precise monitor may provide more consistent and robust associations between sedentary behavior and health outcomes.

The second aim of this study was to determine whether the commonly used AG cut point of 100 counts per minute is the most appropriate cut point for sedentary behavior. Our results suggest the AG150 provided a better estimate of sedentary behavior than the AG100, although there were minimal differences in precision between cut points. Although the AG150 had a smaller estimated bias than the AP, the difference between the two (1.8% and −2.8%, respectively) is small and likely not meaningful. In addition, as discussed above, the AP provides more precise estimates of sedentary behavior than the AG. To determine the source of error in the AP monitor, we examined the difference between AP standing and stepping time. During a 6-h period, 8 min of sitting time was incorrectly classified as standing time, which was overestimated by 11.5 min (stepping time 8 min of sitting time was incorrectly classified as standing time). During a 6-h period, which was overestimated by 11.5 min (stepping time was overestimated by 11.5 min). We did not examine where the error in the AG monitor was because the AG monitor output does not provide standing time.

Recently, Hart et al. (6) examined the convergent validity of the AP and the AG100 and reported that the AG100 resulted in significantly more sedentary time than the AP during a 15-h period, which is not consistent with our results. However, the authors did not report whether the low-frequency extension was used, so it is difficult to interpret the meaning of the magnitude of the differences in sedentary time between studies. Our data were collected with the low-frequency extension filter option selected. The option was added to the GT1M and GT3X monitors by the manufacturer after investigators noted that a greater magnitude of acceleration was required to elicit a nonzero count than was required for the AG 7164 (10,13,18). Therefore, we can only generalize our results to data collected with the 7164 or GT1M/GT3X using the low-frequency extension.

Before selecting a monitor for a study, it is important to consider the purpose of the study and the type of exposure being investigated. On the basis of the results of this study, investigations exclusively focused on the measurement of sedentary behavior should consider using the AP monitor. However, during nonsedentary time, the AP only provides an output of stepping time and cadence of the steps, from which one cannot estimate activity intensity or the type of activity being performed. In contrast, the AG has been used extensively to measure physical activity and exercise time. Using the AG, data processing techniques have been developed to quantify time in MET intensity categories and estimate time in various activity types (e.g., locomotion, sport) (20). Therefore, an individual may consider the AG if a range of activity intensities in addition to or in lieu of sedentary behavior is required.

This study has important limitations that should be noted. First, although DO is considered a criterion measure, human error may affect the accuracy of the DO results. We minimized this by having all observers complete a training program to standardize methods between observers before the commencement of data collection. The AG monitor sampled in 1-s epochs, and it is unlikely the data collector coded the exact second a change in posture occurred. Our study sample was relatively small and included participants who were overweight or obese. We selected this group because approximately 70% of the current US population is overweight, and these individuals are most likely to be targeted for interventions to reduce sedentary behavior (4). It is also important to note that our results may not generalize to individuals whose occupation or lifestyle behaviors included a different set of activities such as an assembly line factory employee or a restaurant worker who stand or are active the majority of the day. Approximately 90% of the observed time was in an office environment where participants were performing employment duties such as computer work, filing papers, delivering messages, and moving around the office building.

There are important strengths to this study. We directly observed participants for more than 1000 h while the monitors were worn. To our knowledge, no other study has validated both the AP and AG monitors in a free-living environment using DO as a criterion measure. An additional strength was that we assessed the monitor’s sensitivity to detect change in behavior by comparing a sedentary condition to an active condition. Activity monitors are commonly used in intervention studies to quantify pre–post changes...
and in epidemiological investigations to distinguish patterns of sedentary behavior. Thus, it is critical to consider the sensitivity of activity monitors to changes in patterns of behavior as a standard practice for validation studies.

This article provides the first-known free-living validation of activity monitors compared with a criterion measure of sedentary behavior. The commonly used AG100 cut point underestimate sitting time to a greater extent than the AG150 compared with DO. Researchers using the AG monitor to estimate sedentary behavior should consider using the count cut point of 150 counts per minute. Compared with DO, the AP monitor provides a precise estimate of sedentary behavior, and the AP is sensitive to reductions of sitting time. The lower bias and higher precision of the AP suggest that the AP is a more appropriate monitor for measuring sedentary time than the AG.

The authors thank the participants in the study and Jeffery Sasaki and Dinesh John, who assisted with data collection and interpretation. The authors also thank Natalia Petruski for assistance with direct observation and the American College of Sports Medicine for providing funding through a doctoral student research grant.

The results of the present study do not constitute endorsement by the American College of Sports Medicine.

REFERENCES