Reliability of Scores From Physical Activity Monitors in Adults With Multiple Sclerosis

Robert W. Motl, Weimo Zhu, Youngsik Park, Edward McAuley, Jennifer A. Scott, and Erin M. Snook
University of Illinois at Urbana-Champaign

We examined the reliability of scores from physical activity monitors in a sample of 193 individuals with multiple sclerosis (MS) who wore a pedometer and an accelerometer for a 7-day period. There were no significant differences among days for the pedometer ($p = .12$) or the accelerometer ($p = .15$) indicating that week and weekend days can be analyzed in a single intra-class correlation (ICC) analytic model. The 7 days of monitoring yielded ICC estimates of .93 for both the pedometer and accelerometer, and a minimum of 3 days yielded a reliability of .80 for both the pedometer and accelerometer. Results indicated that physical activity monitor scores are reliable measures of physical activity for individuals with MS.

Multiple sclerosis (MS) is the most prevalent neurological disease among young and middle-aged adults in the United States (National Multiple Sclerosis Society, 2003), and increasing attention is being directed toward understanding physical activity behavior in this population (Motl, McAuley, & Snook, 2005a). The study of physical activity among adults with MS requires evidence for the validity and reliability of scores from physical activity measures. This is necessary for documenting the frequency and distribution of physical activity, determining the amount of physical activity necessary for managing MS-related symptoms and reducing the risks of secondary diseases (e.g., obesity), identifying the determinants of physical activity, and evaluating the influence of interventions for increasing physical activity in individuals with MS.

Researchers have used a variety of methods for measuring physical activity in those with MS, including self-report surveys and objective devices such as accelerometers and pedometers (Motl et al., 2005a). However, nearly 50% of individuals with MS develop cognitive dysfunction (Bobholz & Rao, 2003), including deficits in short-term memory encoding and retrieval, that might influence the accuracy of recalling physical activity behavior (Ng & Kent-Braun, 1997). Objective devices provide more precise information regarding physical activity within a day and over many days and are relatively unobtrusive and provide minimal interference with...
normal patterns of life (Pearson, Busse, Van Deursen, & Wiles, 2004). Accordingly, objective devices such as accelerometers and pedometers are being used with increasing regularity in individuals with MS (Motl et al., 2005a).

There is emerging evidence that scores from accelerometers and pedometers are valid for quantifying physical activity in those with MS (Busse, Pearson, Van Deursen, & Wiles, 2004; Gosney, Scott, Snook, & Motl, 2007; Ng & Kent-Braun, 1997; Motl, McAuley, Snook, & Scott, 2005b; Motl, Snook, McAuley, & Scott, 2006; Pearson et al., 2004). Indeed, three different methods have been used for establishing the validity of scores from accelerometers and pedometers as measures of physical activity among those with MS; this is consistent with the recommendation by Yun and Ulrich (2002) that the validation process should be undertaken using multiple methods that provided empirical evidence for the adequacy and appropriateness of inferences and decisions from scores on a measurement instrument (cf., Messick, 1995). One method involved the demonstration of hypothesized differences in mean scores from pedometers and accelerometers between individuals with MS and matched controls that did not have MS or any other apparent disease (Busse et al., 2004; Ng & Kent-Braun, 1997; Pearson et al., 2004). Another approach involved a demonstration of the accuracy of pedometers against actual steps taken under controlled laboratory conditions among individuals with MS who were ambulatory without an aide (Motl et al., 2005b). The final approach involved an examination and demonstration of moderate-to-high correlations among scores from pedometers, accelerometers, and self-report surveys among individuals with MS who were ambulatory without an aide (Gosney et al., 2007; Motl et al., 2006).

By comparison, we are unaware of research that has systematically examined the reliability of scores from accelerometers and pedometers for measuring physical activity in those with MS. That is, how many days of monitoring and what type of days are necessary to obtain a stable estimation of physical activity using pedometers and accelerometers in those with MS? Such a line of inquiry has been common in samples of adults without an apparent disease. For example, one previous study involved an examination of the number and type of days necessary to estimate mean pedometer-determined steps per day in 90 adults without an apparent disease (Tudor-Locke et al., 2005). Analyses indicated that 3 days of monitoring was necessary to achieve a reliability of .80 and there was minimal influence of type of day on the reliability estimates (Tudor-Locke et al., 2005).

The number and type of days needed to reach a stable estimation of physical activity behavior may vary by measurement instrument and targeted subpopulation. For example, when assessing habitual activity of 3 or 4-year-old children using heart rate monitors, DuRant et al. (1993) reported adequate reliability with 4.3 days of monitoring; when the same group of children was 5, 6, or 7 years old, adequate reliability was reported with 8.4 days of monitoring with the heart rate monitors. As another example, Matthews et al. (2001) measured physical activity using accelerometers and reported that 3 days of monitoring yielded a stable estimation of activity accounts in adults, but 4 and 7 days of monitoring were necessary for stable estimation of minutes of moderate to vigorous activity and inactivity, respectively. This supports the importance of examining the days needed to establish the stable estimation of physical activity behavior whenever targeting a new subpopulation or using a new physical activity measure.
The present study examined the reliability of scores from a pedometer and an accelerometer, two commonly used objective methods, as measures of physical activity in a sample of adults with MS. We believe that such an examination is essential because of the large difference in physical activity consistently demonstrated in those with MS compared with adults without an apparent disease (Motl et al., 2005a). Such a large difference in physical activity might influence the possible application of research on reliability of pedometers and accelerometers in adults without an apparent disease within those with MS. Indeed, the inactive lifestyles of individuals with MS might require a different number and type of day for generating a reliable estimate of physical activity based on pedometers and accelerometers. Hence, consistent with previous research that included adults without an apparent disease (Tudor-Locke et al., 2005), we sought to examine how many days, and what type of days, are required to reliably estimate physical activity behavior in a sample of adults with MS.

**Method**

**Participants**

During a 7-month period, we recruited a convenience sample of individuals with MS through contact with the three chapters of the National Multiple Sclerosis Society. Recruitment was conducted by contacting the facilitators of local self-help groups and placing advertisements in *MS Connection* quarterly publications. Those interested in participation were asked to contact the research team through either e-mail or telephone. This initial contact was followed up by a telephone call from a member of the research team who described the study and its procedures, answered all questions, conducted a brief screening, and, in the event of meeting inclusion criteria and voluntarily agreeing to participate, recorded the participant’s name and home address for mailing the study materials. The inclusion criteria involved (a) having an established definite diagnosis of MS, (b) being relapse free during the previous 30 days, and (c) being independently ambulatory with minimal assistance (i.e., walking with or without a cane). The diagnosis was based on the participant’s self-reporting of type, duration, and history of MS and clinical criteria used for the diagnosis. Individuals were relapse free during the previous 30 days, as we suspected that physical activity would be substantially reduced during and immediately after such an episode. That is, an exacerbation of the disease (i.e., sudden worsening of MS symptoms) would significantly interfere with an individual’s life (National Multiple Sclerosis Society), including all types of physical activity behavior. Participants were ambulatory with minimal assistance as objective physical activity measurement devices have only been validated with individuals who walk with or without a cane, but not a walker, wheelchair, or scooter. Individuals who used a walker, wheelchair, or scooter were excluded from participation.

There were 255 individuals who inquired about the study, and 221 of those individuals underwent an initial screening. Of the 221 individuals, there were 25 who either did not satisfy our inclusion criteria (n = 23) or did not complete the study materials (n = 2) and were excluded from participation. Three of the 196 individuals did not have 7 complete days (5 week days and 2 weekend days) of pedometer and accelerometer data (i.e., there were missing data from one or more
days). This resulted in a sample of 193 individuals (170 females, 23 males) with MS who participated in this study. Of the 193 participants, 171 self-reported being diagnosed with relapsing-remitting MS, 3 self-reported being diagnosed with primary progressive MS, and 19 self-reported being diagnosed with secondary progressive MS. The mean duration of MS (i.e., time since diagnosis) was 9.0 years ($SD = 7.1$). The sample was primarily Caucasian (93%), married (62%), currently employed (56%), and well educated (26% had some college education and 34% were college graduates) with a median annual household income of greater than $40,000 (65%). The mean age of the sample was 46.2 years ($SD = 9.7$). No data were collected regarding the participants’ gait patterns (e.g., asymmetry) and functional limitations (e.g., stair climbing).

### Instruments

Physical activity was measured by a Yamax SW-200 pedometer (Yamax Corporation, Tokyo, Japan) and the ActiGraph single-axis accelerometer (model 7164 version, Health One Technology, Fort Walton Beach, FL). The Yamax SW-200 pedometer was used in our study because this brand has received the most scientific attention in healthy, nondiseased populations. The Yamax SW-200 pedometer digitally displays step counts. Each pedometer was checked for accuracy using a brief walking test (500 steps at 80 m·min$^{-1}$ on a treadmill), and if the error exceeded 1% (i.e., 5 steps), the pedometer was not used in this study. All pedometers passed the walking test and an additional shake test. Participants recorded the day-end pedometer step counts in a log on a daily basis.

The ActiGraph single-axis accelerometer is a small (2.0 × 1.6 × 0.6 inches) and light weight (1.5 ounces) device, that contains a single, vertical axis piezoelectric bender element that generates an electrical signal proportional to the force acting on it. Acceleration detection ranges in magnitude from 0.05 – 3.2 Gs and the frequency response ranges from 0.25 – 2.5 Hz. Motion outside normal human movements is rejected by a band-pass filter. The positive and negative acceleration signal is digitized by an analog-to-digital converter and numerically integrated over a preprogrammed epoch interval yielding activity counts. Hence, activity counts are a summation of accelerations measured during a cycle period that is established along with start time during an initialization phase. The activity counts represent a quantitative measure of activity over time and are linearly related to the intensity of a participant’s physical activity during a cycle period. The cycle period, or epoch, can range from 1 second to 10 minutes, and in this study was set to be 1 minute. Sixty accelerometers were used, and all of the accelerometers were checked for accuracy by the manufacturer; this was done along with battery replacement by Health One Technology before we collected any data. We further checked for accuracy using our own walking test (15 minutes of walking at 80 m·min$^{-1}$ on a treadmill), and all accelerometers provided an average activity counts per min across the 15 minutes of walking that was within 10% of the mean value accumulated across all 60 accelerometers. All 60 accelerometers were accurate based on manufacturer recommended criteria and our walking test. Participants recorded the exact time the units were worn in a log on a daily basis.
Procedure

The procedure for this study was approved by the Institutional Review Board at the University of Illinois at Urbana-Champaign, and all participants provided written informed consent. The study took place from October of 2003 through June of 2004. After initial telephone contact and voluntary agreement of participation, the informed consent document and a kit containing the pedometer and accelerometer were mailed to each participant. The researchers called to make sure the participants received the package, understood the directions, and signed the informed consent document. The participants were instructed to wear the pedometer and accelerometer on the nondominant hip during the waking hours (except while showering, bathing, and swimming) for a seven-day period beginning on Monday; all participants began on Monday and wore the devices through the following Sunday. Waking hours were defined as the moment upon getting out of bed in the morning through the moment of getting into bed in the evening. The devices were not worn during the night while the participants were asleep. The participants recorded the time that the devices were worn and the day-end steps taken on a log, and compliance with wearing the devices was verified by inspection of the minute-by-minute accelerometer data. On the following Monday after wearing the devices for 7 days, participants returned a signed copy of the informed consent document and the study materials in a prestamped and preaddressed envelop through the mail. Each participant received $20 and a pedometer for participation.

Data Reduction and Analysis

Data from the accelerometers were downloaded into a computer and imported into Microsoft Excel for data processing. The activity counts for each of the seven days were summed. This yielded accelerometer data in activity counts per day. The pedometer data were similarly expressed in steps per day.

The days needed for achieving a targeted reliability level (e.g., .80) can be estimated statistically or determined directly. Two commonly used statistical reliability estimation methods are the Spearman-Brown prophecy formula (Safrit & Wood, 1995) and Generalizability Theory (Morrow, 1989). The Spearman-Brown prophecy formula is a relatively simple method where the days needed for achieving a targeted reliability level can be estimated based on the reliability information with two or more measurement points, whereas Generalizability Theory is a more complex approach in which error resources can be defined in a data collection design. The days needed for achieving a targeted reliability can also be determined directly by computing intra-class correlations, a special case of Generalizability Theory, using all combinations of the number of days and type of the days. We used the direct method because of our interest in both the number and type of days needed to achieve stable estimation of physical activity behavior in individuals with MS.

Our data analytic approach has been used successfully among adults without an apparent disease (Tudor-Locke et al., 2005), thereby affording an opportunity for comparison with previous research. All analyses were conducted using SPSS version 13.0 for Windows (SPSS Inc., Chicago, IL). The research questions, how many and what type of days are necessary to reliably estimate usual physical activity, were
addressed using an initial within-subjects ANOVA and then intra-class correlation (ICC) analyses (Atkinson & Nevill, 1998; Tudor-Locke et al., 2005). We used a one-way (7 days), within-subjects ANOVA and the multivariate $F$-statistic to examine if mean steps per day and activity counts per day differed among the 7 days of monitoring; this was an important precursor for our ICC analyses and illustrated if type of day mattered in generating a reliable estimate of physical activity. The multivariate $F$-statistic was used because it does not assume sphericity and day served as the independent variable in the ANOVA. Effect sizes associated with the $F$-statistic were expressed as partial eta-squared. Effect sizes based on mean differences were expressed as Cohen’s $d$. The family-wise error rate was controlled using the Bonferroni adjustment when test of simple effects were conducted. The ICC analysis was conducted using a two-way model with days and participants as fixed and random effects, respectively. We conducted ICC analyses for all possible combinations of 2, 3, 4, 5, 6, and 7 days using the same combinations of days across all 193 participants to determine the number of days of monitoring to achieve an ICC of 0.80; this is the usual accepted value of a multiple day ICC. The ICC analyses involved a mixed-model framework whereby random effects were assumed for subjects and fixed effects were assumed for days; under conditions of differences across types of days then a random effect would be substituted for days in the ICC analyses.

**Results**

Means (with standard deviations and 95% confidence intervals in parentheses) for steps per day from the pedometer and for activity counts per day from the accelerometer for each day of the week are provided in Table 1. On average, the study participants accumulated 5,902 ($SD = 3,239$) steps per day and 223,522 ($SD = 115,358$) activity counts per day over the 7-day monitoring frame.

The within-subjects ANOVA indicated that there was not a statistically significant difference among days for steps per day from the pedometer, $F(6, 187) = 1.70, p = .12, \eta^2 = .05$. There was a similar lack of difference among days for activity counts per day from the accelerometer, $F(6, 187) = 1.59, p = .15, \eta^2 = .05$. This was an important precursor for our ICC analyses and illustrated that type of day does not matter in generating reliability estimates of physical activity.

The descriptive statistics for the ICCs for all combinations of 2 through 6 days of monitoring are provided in Table 2. The ICC for any 2-day combination ranged from .74 (Combination of Wednesday & Sunday) to .85 (Combination of Tuesday & Wednesday) for steps per day from the pedometer and from .72 (Combination of Thursday & Friday) to .87 (Combination of Monday & Wednesday) for activity counts per day from the accelerometer. Any 3-day combination ranged from .82 (Combination of Monday, Thursday, & Sunday) to .88 (Combination of Monday, Tuesday, & Wednesday) for steps per day from the pedometer and from .82 (Combination of Monday, Thursday, & Sunday) to .89 (Combination of Tuesday, Wednesday, & Thursday) for activity counts per day from the accelerometer. All remaining 4, 5, and 6-day combinations either approached or exceed .90 for both the steps per day from the pedometer and activity counts per day from the accelerometer. The ICC for the entire 7 days of monitoring was .93 for both steps per day from the pedometer and activity counts per day from the accelerometer.
Emerging evidence supports that scores from accelerometers and pedometers are valid for quantifying physical activity in those with MS (Busse et al., 2004; Ng & Kent-Braun, 1997; Motl et al., 2005b; Motl et al., 2006; Pearson et al., 2004). However, no previous study has examined the reliability of scores from accelerometers and pedometers as measures of physical activity in this population. We believe that such an examination is important because of the large difference in physical activity consistently demonstrated in those with MS compared with adults without an apparent disease (Motl et al., 2005a). Such a large difference in physical activity might influence the possible application of research on reliability of pedometers and accelerometers in adults without an apparent disease within those with MS. Indeed, the inactive lifestyles of individuals with MS might require a different number and type of day for generating a reliable estimate of physical activity based

### Table 1  Steps per Day From the Pedometer and Activity Counts per Day From the Accelerometer for Each Day for the Sample of 193 Individuals With Multiple Sclerosis

<table>
<thead>
<tr>
<th>Day</th>
<th>Pedometer (steps per day)</th>
<th>Accelerometer (activity counts per day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td>5,792 (3,754; 5,259 – 6,325)</td>
<td>229,789 (143,305; 209,443 – 250,134)</td>
</tr>
<tr>
<td>Tuesday</td>
<td>5,897 (3,493; 5,401 – 6,393)</td>
<td>224,036 (120,435; 206,937 – 241,135)</td>
</tr>
<tr>
<td>Wednesday</td>
<td>6,204 (3,834; 5,660 – 6,748)</td>
<td>228,404 (126,945; 210,381 – 246,428)</td>
</tr>
<tr>
<td>Thursday</td>
<td>5,940 (3,676; 5,418 – 6,461)</td>
<td>213,819 (117,655; 197,114 – 230,523)</td>
</tr>
<tr>
<td>Friday</td>
<td>6,182 (4,062; 5,605 – 6,759)</td>
<td>229,154 (153,194; 207,404 – 250,904)</td>
</tr>
<tr>
<td>Saturday</td>
<td>5,723 (4,237; 5,121 – 6,325)</td>
<td>225,767 (148,413; 204,696 – 246,838)</td>
</tr>
<tr>
<td>Sunday</td>
<td>5,575 (3,918; 5,019 – 6,131)</td>
<td>213,683 (143,305; 193,337 – 234,029)</td>
</tr>
<tr>
<td>Week (All days)</td>
<td>5,902 (3,239; 5,442 – 6,362)</td>
<td>223,522 (115,358; 207,144 – 239,900)</td>
</tr>
</tbody>
</table>

Note. Values are mean (standard deviation; 95% confidence interval).

### Table 2  Descriptive Statistics for the Intra-Class Correlations for the Pedometer and the Accelerometer Across Combinations of 2 – 6 Days of Monitoring for the Sample of 193 Individuals With Multiple Sclerosis

<table>
<thead>
<tr>
<th>Combinations of days</th>
<th>Pedometer</th>
<th>Accelerometer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>2 Days</td>
<td>.79</td>
<td>.79</td>
</tr>
<tr>
<td>3 Days</td>
<td>.85</td>
<td>.85</td>
</tr>
<tr>
<td>4 Days</td>
<td>.88</td>
<td>.88</td>
</tr>
<tr>
<td>5 Days</td>
<td>.90</td>
<td>.90</td>
</tr>
<tr>
<td>6 Days</td>
<td>.92</td>
<td>.92</td>
</tr>
</tbody>
</table>

**Discussion**

Emerging evidence supports that scores from accelerometers and pedometers are valid for quantifying physical activity in those with MS (Busse et al., 2004; Ng & Kent-Braun, 1997; Motl et al., 2005b; Motl et al., 2006; Pearson et al., 2004). However, no previous study has examined the reliability of scores from accelerometers and pedometers as measures of physical activity in this population. We believe that such an examination is important because of the large difference in physical activity consistently demonstrated in those with MS compared with adults without an apparent disease (Motl et al., 2005a). Such a large difference in physical activity might influence the possible application of research on reliability of pedometers and accelerometers in adults without an apparent disease within those with MS. Indeed, the inactive lifestyles of individuals with MS might require a different number and type of day for generating a reliable estimate of physical activity based
on pedometers and accelerometers. Our results suggest that a minimum of 3 days of pedometer and accelerometer data are sufficient for reliably estimating physical activity among adults with MS, and the type of day (e.g., weekday or weekend day) does not influence the reliability estimates.

Our findings regarding reliability are in agreement with previous research using adults without an apparent disease (Tudor-Locke et al., 2005). That previous study involved an examination of the number and type of days necessary to estimate mean pedometer-determined steps per day in 90 healthy adults who did not have an apparent disease using the Yamax SW-200 pedometer. Analyses indicated that 3 days of monitoring was necessary to achieve a reliability of .80, and there was little influence of type of day on the reliability estimates. Using both pedometers and accelerometers, we similarly report that 3 days of monitoring was necessary to achieve a reliability of .80, and there was little influence of type of day on the reliability estimates. This convergence between distinct samples confirms the recommendation that any 3 days of monitoring can provide a sufficiently reliable estimate of physical activity (Tudor-Locke et al., 2005). This seems to be true for regularly active adults and inactive adults who have a neurological disease such as MS.

Our results have important design implications for future research among those with MS, particularly for guiding assessment protocols, reducing measurement costs and respondent burden, and enhancing participant compliance. That is, the minimum of 3 days of monitoring, compared with 7 days, will reduce participant burden and likely enhance compliance with wearing the devices. Importantly, the requirement of only 3 days of monitoring will reduce the overall costs of measuring physical activity using pedometers and accelerometers. That is, shorter measurement periods will require fewer devices to measure physical activity in a large group of participants over the same period of time.

The strengths of this study include a novel focus on reliability of scores from 2 objective devices as measures of physical activity in a large sample of individuals with MS. Importantly this study is not without limitations. One limitation is that we only focused on the number and type of days of pedometer and accelerometer monitoring that represented a week of physical activity in those with MS. Future researchers might consider addressing how many days or weeks of monitoring represent a month or year of physical activity in those with MS. Another limitation is the demographic characteristics of the participants. Although the descriptive features of our sample are roughly consistent with the demographics of MS, the sample primarily consisted of women with relapsing-remitting MS and who were Caucasian. Future researchers might consider using a more diverse sample that included a larger number of men, individuals with primary and secondary progressive MS, and minorities. One final limitation is the possible problem associated with compound symmetry in examinations of reliability, and this problem was unaccounted for in the present investigation.

In summary, we report that 3 days of monitoring using both pedometers and accelerometers was necessary to achieve a reliability of .80 in a large sample of individuals with MS, and there was little influence of type of day on the reliability estimates. Our results are strikingly similar with previous research using healthy nondisease adults (Tudor-Locke et al., 2005), and such convergence between distinct samples further supports the recommendation that any 3 days of monitoring can provide a sufficiently reliable estimate of physical activity (Tudor-Locke et al., 2005).
References


