Reliability and Validity of Bilateral Ankle Accelerometer Algorithms for Activity Recognition and Walking Speed After Stroke

Bruce H. Dobkin, MD; Xiaoyu Xu, BS; Maxim Batalin, PhD; Seth Thomas, BS; William Kaiser, PhD

Background and Purpose—Outcome measures of mobility for large stroke trials are limited to timed walks for short distances in a laboratory, step counters and ordinal scales of disability and quality of life. Continuous monitoring and outcome measurements of the type and quantity of activity in the community would provide direct data about daily performance, including compliance with exercise and skills practice during routine care and clinical trials.

Methods—Twelve adults with impaired ambulation from hemiparetic stroke and 6 healthy controls wore triaxial accelerometers on their ankles. Walking speed for repeated outdoor walks was determined by machine-learning algorithms and compared to a stopwatch calculation of speed for distances not known to the algorithm. The reliability of recognizing walking, exercise, and cycling by the algorithms was compared to activity logs.

Results—A high correlation was found between stopwatch-measured outdoor walking speed and algorithm-calculated speed (Pearson coefficient, 0.98; \( P = 0.001 \)) and for repeated measures of algorithm-derived walking speed (\( P = 0.01 \)). Bouts of walking >5 steps, variations in walking speed, cycling, stair climbing, and leg exercises were correctly identified during a day in the community. Compared to healthy subjects, those with stroke were, as expected, more sedentary and slower, and their gait revealed high paretic-to-unaffected leg swing ratios.

Conclusions—Test–retest reliability and concurrent and construct validity are high for activity pattern-recognition Bayesian algorithms developed from inertial sensors. This ratio scale data can provide real-world monitoring and outcome measurements of lower extremity activities and walking speed for stroke and rehabilitation studies. (Stroke. 2011;42:2246-2250.)

Key Words: accelerometry ▪ activity measures ▪ ambulatory monitor ▪ stroke outcomes ▪ walking speed

Randomized clinical trials to lessen physical impairments and disabilities with interventions for acute stroke and for rehabilitation usually include surrogate measurements of real-world, mobility-related activities. For example, the laboratory-based 10-m walking speed and the distance walked in 2 to 6 minutes are often used to assess walking ability and capacity. These tools are supplemented by ordinal disability scales, such as the Modified Rankin Scale and Functional Independence Measure.\(^1–6\) Quality-of-life questionnaires, such as the Stroke Impact Scale,\(^7\) add a subject’s perception about daily activity and participation. These are indirect measures in that they do not quantify activity or walking speed and distance as actually performed outside the laboratory. Levels of exercise and mobility are remarkably low during rehabilitation and in daily life after stroke, and are associated with poor cardiovascular fitness.\(^8\) Efforts to augment physical training and skills learning to optimize walking speed,\(^9\) balance, distances walked, and fitness could improve daily functioning, reduce risk factors for recurrent stroke,\(^10\) and possibly improve aspects of cognition that often decline with stroke and aging.\(^11–12\)

One confounding problem in the development of such interventions has been the absence of ratio scale tools to continuously identify in patients the type, quantity, and aspects of the quality of practice and activities outside of the clinic.

Direct measurements of activity and mobility in the home and community would offer greater ecological validity about the efficacy of poststroke interventions,\(^13\) as well as provide feedback about practice and activity to encourage compliance. For mobility-related outcomes, inexpensive single accelerometers that serve as movement counters have been deployed to reveal the number of steps taken outside of the laboratory, but these devices are limited in their accuracy and cannot assess walking speed, distance, gait variations under changing environmental conditions, or describe asymmetries between the hemiparetic and less affected leg as a measure of quality of gait.\(^9,14\)
The cost of more sophisticated sensor technologies, such as multiple triaxial accelerometers, is rapidly decreasing. Algorithms enable raw inertial data to provide more specific information, such as walking speed, stair climbing, and leg exercise. Using newly developed mathematical algorithms with machine-learning computations, we tested the reliability of wireless triaxial accelerometers placed on each ankle of healthy subjects and those with hemiparesis after stroke to accurately detect all bouts of walking, along with walking speed and lower extremity exercise in the community.

Subjects and Methods
A convenience sample was recruited from the Southern California Stroke Association. Participants met entry criteria of hemiparesis, having no better than movement against light resistance at the hip flexors and knee extensors and independent walking in the home. Exclusion criteria included lower extremity contractures or pain, cardiopulmonary disease that limited exercise tolerance, and recent medical complications. Descriptors appear in Table 1. The healthy controls were a convenience sample of 3 males and 3 females aged 30 to 60 years old (mean, 40 years). The study was approved by the Institutional Review Board.

Accelerometry Procedures
Wireless triaxial accelerometers, attached by Velcro (3M Corporation, St Paul, MN) to a soft snap band, were placed above each ankle on the bony tibia, 3 cm above the medial malleous, with the lateral edge flush against the anterior border of the tibia (Figure) to prevent displacement. The y-axis is vertical. We used low-cost ($130), 100-g, 1 g, 1×2×5-cm annular devices (Gulf Coast Data Concepts) with a USB port for charging and downloading. The sampling rate was 320 Hz.

Subjects performed stopwatch-timed 50-ft walks at slow, casual, and fastest speeds on a flat indoor surface. All walks began with 1 foot behind the other that was on the start line. After saying “ready, go,” the stopwatch started when the hind foot first crossed the start line, and then stopped when the lead foot first crossed the end line. Subjects used usual assistive devices. This test of walking speed served as a template for the initial activity pattern-recognition algorithm of the raw sensor data for each subject (Table 2). The subjects then walked from the clinic to the outdoors. For subjects 2 to 9, this ∼300-ft untimed walk (Supplemental Table I, http://stroke.ahajournals.org) included crossing a 4-lane (62-ft-wide) street with a traffic light. On a sidewalk with pedestrians and automobile traffic 15 ft on 1 side, subjects walked at their usual speed for 67 ft twice, and then continuously with a turn on the 67-ft course (134 ft). Each segment was timed by a stopwatch and served as the ground truth for comparison with subsequent calculations by the algorithms. Subjects wore the sensors home for the next 24 hours when out of bed. They kept a log of mobility activities. This checklist included walking in the home, community, or at work, and kitchen tasks, dressing, sitting, and exercising.

The healthy control subjects performed three 50-ft timed walks at slow to fast speeds and ascended and descended 5 stairs for the baseline template of activity. They also wore ankle sensors and kept a log for 24 hours of activity.

Statistics
The sensor data were analyzed by machine-learning algorithms with no knowledge of the distance walked for the timed 67-ft outdoor test. A blinded observer compared a printout of the bins of activity derived from the algorithm to the time slots of activity in each participant’s log. For the primary reliability experiment of repeated measures of outdoor walking speed and for the primary concurrent validity study of stopwatch-timed versus algorithm-timed outdoor 67-ft walks, we used the Pearson correlation coefficient. The error SD was the SD of the differences between the methods. The bias was the mean of these differences. The mean absolute deviation was computed by taking the differences between the methods for each observation and then computing the absolute value and the mean. The same statistic was used for comparing observer counts of steps taken to accelerometer step counts during the transition from testing indoors to walking to the outdoor 67-ft course in subjects 2 to 9.

Sensor Analysis System
The Medical Daily Activity Wireless Network (MDAWN) was designed as a complete architecture to identify and quantify purposeful leg movements. The accelerometer signal processing and activity state classification system includes components for automated sensor data collection, transport to a secure remote repository, individualized subject model development, and classification by sensor fusion analysis principles. The system is hosted at the MDAWN DataServer at University of California Los Angeles. The naïve Bayes classifier relies on a probabilistic model of inertial data features, classes, and their relation. The feature extraction step for
Table 2. Range of Walking Speeds of Subjects With Stroke Who Were Instructed to Walk at Fast, Usual, and Slow Speeds

<table>
<thead>
<tr>
<th>Speed</th>
<th>Subject 1</th>
<th>Subject 2</th>
<th>Subject 3</th>
<th>Subject 4</th>
<th>Subject 5</th>
<th>Subject 6</th>
<th>Subject 7</th>
<th>Subject 8</th>
<th>Subject 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>50-ft Indoor Speed</td>
<td>0.47 F</td>
<td>0.59 F</td>
<td>0.57 F</td>
<td>1.15 F</td>
<td>0.76 F</td>
<td>0.57 F</td>
<td>0.54 F</td>
<td>0.48 U</td>
<td>0.41 S</td>
</tr>
<tr>
<td>0.46 U</td>
<td>0.55 U</td>
<td>0.58 U</td>
<td>1.21 U</td>
<td>0.88 U</td>
<td>0.42 U</td>
<td>0.60 U</td>
<td>0.40 U</td>
<td>0.51 S</td>
<td>0.53 U</td>
</tr>
<tr>
<td>0.41 S</td>
<td>0.60 S</td>
<td>0.54 U</td>
<td>0.62 U</td>
<td>1.22 U</td>
<td>0.62 U</td>
<td>0.61 U</td>
<td>0.40 U</td>
<td>0.50 S</td>
<td>0.50 U</td>
</tr>
<tr>
<td>0.58 U</td>
<td>0.57 S</td>
<td>0.53 U</td>
<td>0.53 U</td>
<td>1.15 U</td>
<td>0.54 U</td>
<td>0.53 U</td>
<td>0.39 U</td>
<td>0.50 U</td>
<td>0.59 U</td>
</tr>
<tr>
<td>0.51 S</td>
<td>0.54 U</td>
<td>0.50 U</td>
<td>0.50 U</td>
<td>1.70 U</td>
<td>0.57 U</td>
<td>0.50 U</td>
<td>0.39 U</td>
<td>0.50 U</td>
<td>0.59 U</td>
</tr>
<tr>
<td>0.50 S</td>
<td>0.48 F</td>
<td>0.51 F</td>
<td>0.51 F</td>
<td>1.25 S</td>
<td>0.57 F</td>
<td>0.50 F</td>
<td>0.40 F</td>
<td>0.48 F</td>
<td>0.50 F</td>
</tr>
<tr>
<td>0.49 S</td>
<td>0.54 S</td>
<td>0.53 S</td>
<td>0.53 S</td>
<td>1.15 S</td>
<td>0.57 S</td>
<td>0.53 S</td>
<td>0.39 S</td>
<td>0.48 S</td>
<td>0.50 S</td>
</tr>
<tr>
<td>0.47 S</td>
<td>0.54 S</td>
<td>0.50 S</td>
<td>0.50 S</td>
<td>1.25 S</td>
<td>0.54 S</td>
<td>0.50 S</td>
<td>0.39 S</td>
<td>0.47 S</td>
<td>0.50 S</td>
</tr>
</tbody>
</table>

Outdoor walking speed was calculated by stopwatch (ground truth) vs Medical Daily Activity Wireless Network algorithm for up to 3 bouts of ambulation.

F indicates fast; S, slow; U, usual.

Walking Speed

The fast, usual, and slow 50-ft indoor walking speeds for hemiparetic subjects are shown in Table 2 (second column). Based on the algorithm derived from this baseline sensor data, subsequent walking activity and speeds were calculated for the outdoor timed walks by MDAWN. Table 2 shows the primary concurrent validity comparison between stopwatch-timed and algorithm-derived outdoor walking speeds, as well as their relationship to the indoor range of walking speeds. The Pearson correlation between stopwatch-measured outdoor walking speed and algorithm-calculated walking speed was 0.98 (P=0.001), with a mean absolute deviation of 0.045 (6.7% of the average speed of 0.67), a bias of 0.01, and an error SD of 0.05. No outliers were found. Concurrent validity, then, was high (P=0.001). In addition, the test–retest reliability for these walks performed 3 times by subjects 1 to 9 was high (P=0.01), with no outliers (Table 2). In retrospect, subjects who walked >0.8 m/sec (unlimited community walkers)23 showed the widest range of indoor walking speeds from slow to fast. Walking speeds crossing the 62-ft-wide pedestrian street for 2 of 6 subjects who were timed (not shown) were 11% and 15% faster than their fastest walking speed timed indoors, 2 were within 5%, and 2 were 10% and 16% slower.

For the 6 healthy subjects (data not shown), the Pearson correlation was 0.98 (P=0.001), revealing concurrent validity of outdoor speed measures. The significant difference in mean walking speeds between healthy subjects and those with stroke supports the construct validity of the algorithms.

Walking Activity

During the transition from the indoor to outdoor walkway, the steps taken by subjects 2 to 9 were counted by an investigator and, later, calculated by the algorithm, along with walking speed. The step counts highly correlated (0.99) with each other, with no outliers. Walking speeds are shown in Supplemental Table I (Transition to Outdoors column). These speeds were similar to the usual walking speed during the 50-ft timed walk for most subjects, providing additional construct validity. All subjects then wore their accelerometers for 1 day. Supplemental Table I shows walking during home and community activity. Bouts of walking were defined as a minimum of 5 strides and identified by the activity pattern-recognition algorithm. Speeds varied over the day within subjects but were within the range of the slow-to-fast 50-ft indoor velocities (Table 2). The time spent walking, mean duration of bouts (28–77 seconds), and number of steps taken (383–5880) varied across subjects. Subject 10, the fastest walker, was the only person who achieved the mean number of steps taken daily by our healthy subjects (>9000). The
construct validity of the algorithms was also revealed by the uneven ratio of swing time of the affected lower extremity compared to the unaffected one, which ranged from 1.1 to 1.7. Healthy subjects had normal ratios for the right leg versus left leg of 0.95 to 1.05. Thus, this aspect of quality of movement was identified.

**Activity Detection**

The pattern-recognition algorithms were annotated for the time and type of each block of activity. The logs were then compared for defined categories of mobility and exercise. For the subjects with stroke, frequency profiles of walking obtained from the sensors all took place within the time frame of documented mobility-related activities. In addition, stationary bicycling, alternating leg lifts, and repetitive heel slides while supine on a mat were identified by MDAWN. All periods of mobility and lower extremity exercise were correctly categorized in healthy people as well, including walking and stationary bicycling at various speeds, jogging, and stair climbing. Some repetitive leg movements could not be classified at first. Sensor data were re-examined in relation to the activity logs. Once the algorithm was given a definition of the discrete classification of movement it had identified, it correctly identified further bouts of elliptical machine exercise and resistance exercise of the leg against an elastic band. The findings support convergent validity for the algorithms and logs.

**Discussion**

Using wireless ankle accelerometers, the MDAWN algorithms identified walking speeds with high concurrent or criterion validity (stopwatch versus algorithm) and test–retest reliability (reproducibility of speed measures at different times). MDAWN activity pattern recognition also had high convergent or construct validity compared to logs in identifying bouts of walking, cycling, and leg exercises. The results held for disabled persons with hemiparesis and healthy subjects. Cadence and distance walked also can be derived from these data. We found additional construct validity in that the speeds measured indoors and outdoors were similar within subjects regarding the velocities used in the home and community. In addition, participants with stroke, as expected, walked slower, had shorter bouts of walking, and had higher swing ratios compared to our convenience sample of healthy subjects. As anticipated, subjects with stroke, especially those who walked at slower speeds, were remarkably inactive over 24 hours, as has been found by others. The face validity of the measures derived from MDAWN seems high as well, because ecologically sound assessments of real-world, mobility-related activity have been a high priority for rehabilitation. The results also confirm reports that the laboratory-based walking speed may not replicate the habitual walking speed of hemiplegic subjects when they move about in the community.

The algorithms provided insight into an aspect of quality of gait by revealing the duration of swing times for each leg under different conditions. Collectible parameters relevant to the quantity of movement included repetitions per minute, duration of bouts of activity, walking distance and cadence, and ratio measures to permit calculation of means and quartiles for this data. Thus, monitoring with the MDAWN system provided generally unobtainable information about poststroke daily mobility and exercise activities in the home and community, despite the use of ankle–foot orthoses, canes, and patterns of walking that differ markedly from that of healthy persons.

Accelerometry has made inroads in subacute stroke and rehabilitation outpatient studies to count steps of 1 leg, but not for walking speed. Other algorithms have distinguished a limited set of activities or identified the step cycle. The machine-learning algorithms of the MDAWN system overcame some of the limitations of commercial monitors. These systems use proprietary analytics that limit their flexibility for research and the sensors are too expensive and fragile to distribute to participants in clinical trials. Available systems do measure cadence, step time, and speed, and have high repeatability (intraclass correlation >0.9) and interobserver reliability (intraclass correlation >0.8) across all ages and genders, unless accelerometer signals are collected at <180 Hz.

The results move forward the ability of clinical trialists to make reliable measurements of the type, quantity, and aspects of quality of physical activities. New tools like MDAWN can augment ordinal scales by assessing not only the perception or capacity for an activity but also what patients actually perform daily. Patients after stroke often overestimate their level of daily activity.

Although this pilot study begins to establish the reliability and validity of sensor activity recognition and quantification, we did not attempt an in-depth analysis of this convenience sample. To fully capture the typical amount of daily ambulation and range of activities, 5 to 7 days of acquisition are necessary. An additional sensor attached to the chest or thigh, as well as global positioning satellite and gyroscope data, can add capabilities as needed. Algorithms can fuse these data.

Using inertial sensors with recognition algorithms, future studies could remotely monitor the fidelity and quantity of skills practice after stroke for rehabilitation trials and compliance with fitness recommendations. Remote monitoring may allow more patients to participate in trials at low costs. Continuous recordings may reveal the trajectory of gains in lower extremity activities, enable dose–response curves for interventions, and detect declines in activity as an early warning of disease exacerbation.

**Conclusions**

The MDAWN system enables the direct evaluation of walking and other physical activities of patients in the community, rather than evaluation solely by surrogate laboratory measurements and questionnaires. This tool may offer inexpensive, quantifiable, clinically meaningful monitoring and outcome data for randomized clinical trials and individual care.

**Sources of Funding**

This material is partly based upon work supported by the National Science Foundation under Grant No. 0120778 (The Center for Embedded Networked Systems). Any opinions, findings, and con-
clusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Disclosures
None.

References
Reliability and Validity of Bilateral Ankle Accelerometer Algorithms for Activity Recognition and Walking Speed After Stroke

Bruce H. Dobkin, Xiaoyu Xu, Maxim Batalin, Seth Thomas and William Kaiser

Stroke. 2011;42:2246-2250; originally published online June 2, 2011;
doi: 10.1161/STROKEAHA.110.611095

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2011 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/42/8/2246

Data Supplement (unedited) at:
http://stroke.ahajournals.org/content/suppl/2011/06/02/STROKEAHA.110.611095.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at:
http://stroke.ahajournals.org//subscriptions/
Reliability and validity of bilateral ankle accelerometer algorithms for activity recognition and walking speed after stroke

Supplementary Table A.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Transition to Outdoors</th>
<th>Home and Community Walking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Speed Range m/s</td>
<td>Avg m/s</td>
</tr>
<tr>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>2</td>
<td>0.54 - 0.57</td>
<td>0.56</td>
</tr>
<tr>
<td>3</td>
<td>0.50 - 0.53</td>
<td>0.51</td>
</tr>
<tr>
<td>4</td>
<td>0.85 – 0.89</td>
<td>0.87</td>
</tr>
<tr>
<td>5</td>
<td>0.80 - 0.86</td>
<td>0.83</td>
</tr>
<tr>
<td>6</td>
<td>0.52 - 0.57</td>
<td>0.54</td>
</tr>
<tr>
<td>7</td>
<td>0.15 - 0.40</td>
<td>0.36</td>
</tr>
<tr>
<td>8</td>
<td>0.78 - 0.80</td>
<td>0.79</td>
</tr>
<tr>
<td>9</td>
<td>0.84 - 1.02</td>
<td>0.95</td>
</tr>
<tr>
<td>10</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>11</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>12</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Speed calculated by the MDAWN algorithm that was used to walk to the outdoor test site (Transition) by stroke participants. Speed and exemplars of calculated parameters (range of speeds used, number and duration of bouts of walking) are shown in remaining columns for one day of community activity.