MSJ Experimental Translational Clinical

Original Article

Accuracy and precision of smartphone applications and commercially available motion sensors in multiple sclerosis

Julia M Balto, Dominique L Kinnett-Hopkins and Robert W Motl

Abstract

Background: There is increased interest in the application of smartphone applications and wearable motion sensors among multiple sclerosis (MS) patients.

Objective: This study examined the accuracy and precision of common smartphone applications and motion sensors for measuring steps taken by MS patients while walking on a treadmill.

Methods: Forty-five MS patients (Expanded Disability Status Scale (EDSS) = 1.0–5.0) underwent two 500-step walking trials at comfortable walking speed on a treadmill. Participants wore five motion sensors: the Digi-Walker SW-200 pedometer (Yamax), the UP2 and UP Move (Jawbone), and the Flex and One (Fitbit). The smartphone applications were Health (Apple), Health Mate (Withings), and Moves (ProtoGeo Oy).

Results: The Fitbit One had the best absolute (mean = 490.6 steps, 95% confidence interval (CI) = 485.6–495.5 steps) and relative accuracy (1.9% error), and absolute (SD = 16.4) and relative precision (coefficient of variation (CV) = 0.0), for the first 500-step walking trial; this was repeated with the second trial. Relative accuracy was correlated with slower walking speed for the first ($r_s = -.53$) and second ($r_s = -.53$) trials.

Conclusion: The results suggest that the waist-worn Fitbit One is the most precise and accurate sensor for measuring steps when walking on a treadmill, but future research is needed (testing the device across a broader range of disability, at different speeds, and in real-life walking conditions) before inclusion in clinical research and practice with MS patients.

Keywords: Motion sensors, Fitbit, smartphone applications, steps, MS

Introduction

There is increasing interest in the adoption of smartphone applications and wearable motion sensors (e.g. Digi-Walker SW-200 (Yamax), One (Fitbit), and UP Move (Jawbone) for measuring community-based, ambulatory physical activity in clinical research and practice involving persons with multiple sclerosis (MS).[1] For example, researchers from Biogen Idec recently teamed with PatientsLikeMe and examined the feasibility of using a consumer-wearable motion sensor for monitoring activity among people with MS in a real-world setting.[2] Two hundred forty-eight PatientsLikeMe members were provided with the waist-worn Fitbit One device, and 82% (213) activated it and authorized data access by PatientsLikeMe. Of those persons, 95% (203) synchronized the device and produced tracking data. Participants synced the devices on an average of 18.2 days over the 21-day study period and walked an average of 4671 steps per day. Further post-study survey data were available from 191 participants, and 88% of respondents reported that the device was easy to use and incorporate into daily life; 83% reported interest in continued use of the device; and 68% believed that the device would be useful for self-managing MS. These feasibility data provide a strong case for the usability of smartphone applications and wearable motion sensors in clinical research and practice for measuring community-based, ambulatory physical activity in MS.

Correspondence to:
Robert W Motl
906 S. Goodwin Ave, Freer
Hall, Urbana IL, 61801, USA
robmotl@illinois.edu

**Julia M Balto, Dominique
L Kinnett-Hopkins, Robert
W Motl**
Department of Kinesiology
and Community Health,
University of Illinois at
Urbana-Champaign, IL, USA

The next step in this adoption process requires research examining the accuracy and precision of applications and motion sensors for measuring a common outcome of steps taken while ambulating using a standard protocol. To date, we are unaware of published research evaluating the accuracy and precision of smartphone applications and wearable motion sensors for measuring actual steps taken while ambulating in persons with MS, although these properties have been examined in the general population.[3] Those researchers examined the accuracy and precision of 10 applications and motion sensors for measuring actual steps taken during trials of walking 500 and 1500 steps on a treadmill in a sample of 14 healthy adults. The relative error (i.e. accuracy) in measuring actual steps taken ranged between −0.3% and 1.0% for the waist-worn sensors (i.e. Digi-Walker SW-200 pedometer (Yamax), Zip (Fitbit), and One (Fitbit)), between −22.7% and −1.5% for the wrist-worn motion sensors (i.e. Flex (Fitbit), UP24 (Jawbone), and Fuelband (Nike)), and between −6.7% and 6.2% for smartphone applications (i.e., Fitbit (Fitbit), Health Mate (Withings), and Moves (ProtoGeo Oy)). The Fuelband (Nike) had the greatest amount of variance (i.e. worst precision), whereas the One (Fitbit) and Zip (Fitbit) had the least variance (i.e. best precision).

This study adopted a similar protocol and examined the accuracy and precision of three smartphone applications and five wearable motion sensors for measuring actual steps taken, using direct observation as a gold standard, while walking 500 steps on a treadmill, in a sample of persons with MS. Such examination is necessary, as persons with MS walk more slowly and have gait alterations compared with adults from the general population, and such patterns might influence the accuracy and precision of the applications and motion sensors for measuring steps taken while walking.[4,5]

Methods

Participants

Community-dwelling persons with MS who had participated in previous laboratory studies were invited through telephone calls and email messages to participate in this study. Inclusion criteria for participation were (a) neurologist-confirmed diagnosis of MS; (b) age between 18 and 64 years; (c) relapse free for the past 30 days; and (d) ability to walk 500 steps without using an assistive device. A total of 57 individuals were contacted, and 11 were uninterested in participating. The resulting 46 persons underwent screening, and 45 met inclusion criteria and were

Table 1. Demographic and clinical characteristics.

	Mean	SD	Range
Age (years)	46.7	10.0	23.0–62.0
Height (cm)	169.4	10.7	154.0–204.2
Weight (kg)	76.2	18.4	48.2–131.5
BMI (kg/m^2)	26.6	6.1	18.7–44.5
MS duration (years)	11.4	9.3	0.0–35.0
EDSS (median, IQR)	3.0	1.5	1.0–5.0
MS type (% RRMS)	95.60	—	—
MSWS-12	18.5	19.9	0.0–77.1
Walking speed (mph)	2.7	0.6	1.5–4.0

BMI: body mass index; IQR: interquartile range; SD: standard deviation.

scheduled for testing. The final sample included 45 persons with MS. The demographics and clinical characteristics of the participants are provided in Table 1.

Physical activity trackers

This study included five wearable motion sensors: Digi-Walker SW-200 pedometer (Yamax), UP2 (Jawbone), UP Move (Jawbone), Flex (Fitbit), and One (Fitbit). The study further included three physical activity smartphone applications: Health (Apple), Health Mate (Withings), and Moves (ProtoGeo Oy). We selected these motion sensors and applications based on evaluation in previous research in the general population,[3] popularity, and availability at the time of investigation. Details about these motion sensors and applications are provided in Table 2. The Digi-Walker pedometer was placed in a position on the participant's non-dominant hip that optimized the accuracy of counts based on the 20-step test. The One and UP Move were randomly placed on either side of the Digi-Walker. The UP2 (Jawbone) and the Flex (Fitbit) were placed on the wrist of the participant's non-dominant wrist in random order. An Apple iPhone 5 was placed in the participant's front non-dominant pocket with the Health (Apple), Health Mate (Withings), and Moves (ProtoGeo Oy) applications running concurrently in the background.

Disability and walking status

All participants underwent a neurological examination by a Neurostatus certified examiner for generation of an Expanded Disability Status Score (EDSS) score for describing the sample.[4] Participants further completed a self-report measure of mobility. The Multiple Sclerosis Walking Scale (MSWS-12) is a 12-item, patient-rated measure of walking

Table 2. Motion sensor and smartphone application characteristics.

Tracker	Sensor type	Wear site	Data interface and version	Setup parameters	Retail price
Digi-Walker SW-200 pedometer (Yamax)	Spring loaded	Waist (non-dominant side)	Highly readable LCD display on device	–	$20
UP2 (Jawbone)	Triaxial accelerometer	Wrist (non-dominant side)	Jawbone UP V.4.4.1	H, W, sex, DOB	$99.99
UP Move (Jawbone)	Triaxial accelerometer	Waist (non-dominant side)	Jawbone Move V.4.4.1	H, W, sex, DOB	$49.99
Flex (Fitbit)	Triaxial accelerometer	Wrist (non-dominant side)	Fitbit, 2.10	H, W, sex, DOB	$99.95
One (Fitbit)	Triaxial accelerometer	Waist (non-dominant side)	Fitbit, 2.10	H, W, sex, DOB	$99.95
Health (Apple)	M7 motion coprocessor	Phone, front pocket, non-dominant side	Apple, iPhone 6	H, W, DOB	Free (with phone)
Health Mate (Withings)	M7 motion coprocessor	Phone, front pocket, non-dominant side	Withings Health Mate V.2.6.0	H, W, sex, DOB	Free (with phone)
Moves (ProtoGeo Oy)	M7 motion coprocessor	Phone, front pocket, non-dominant side	Moves ProtoGeo V.2.5	H, W, sex, DOB	Free (with phone)

DOB: date of birth; H, height; W, weight.

impairment that has been validated in persons with MS.[5] The overall MSWS-12 score ranges between 0 and 100, and higher scores reflect worse perceived ambulation.

Usual walking speed
Participants completed four over-ground trials of the timed 25-foot walk (T25FW). Participants were instructed to walk at a comfortable, normal pace during the four T25FW trials that mimicked ambulation during daily life. The primary outcome of the T25FW trials was speed (mph). T25FW speed was averaged across the four trials, and this generated an estimate of comfortable walking speed for the subsequent treadmill protocol. To that end, this overground protocol provided an estimate of comfortable walking speed and permitted a safe assessment of device accuracy under a controlled walking speed condition that approximated conditions of real-world ambulation.

Treadmill protocol
The accuracy and precision of the motion sensors and smartphone applications were evaluated against the actual steps taken, recorded through direct observation, while walking on a motor-driven treadmill (Trackmaster model TMX425C, Full Vision, Inc., Newston, KS). We provided a 5-min rest period following the T25FW trials, while a researcher fitted the participant with all of the wearable motion sensors and placed the iPhone 5 in the participant's front pocket. The participants began the treadmill walks with an acclimation trial. This involved walking on the treadmill at a comfortable pace (determined by the average speed of the preceding T25FW trials) for a minimum of 1 min and maximum of 5 min; the acclimation trial stopped when the participant reported feeling comfortable with walking on the treadmill. Participants then rested for 5 min by sitting quietly in a comfortable chair placed on the stopped treadmill track. Participants next walked for 500 steps on the treadmill at the same speed as the acclimation trial (i.e. comfortable walking speed). The actual number of steps taken by the participant was measured in duplicate by two researchers using hand-tally counters, and the total steps taken and those measured by the sensors and applications were recorded in a data log. To ensure accuracy of the observed steps, researchers compared step counts every 20 steps and restarted the trial if the hand-tally-counted steps differed; this never occurred. Participants again rested for 5 min on the comfortable chair and then undertook a second 500-step trial for repeatability.

Procedures
The procedures for this study were approved by the University of Illinois at Urbana-Champaign

Institutional Review Board, and all participants provided written informed consent. Participants initially completed a demographic questionnaire and the MSWS-12. Participants underwent neurological examination, measurement of height and weight using a scale stadiometer (Detecto, Webb City Missouri), and assessment of usual walking speed. The participants then undertook the treadmill protocol, and were remunerated $15.

Statistical analysis

We expressed accuracy and precision based on absolute and relative metrics. Absolute accuracy involved the mean and 95% confidence interval (CI) for actual steps recorded per device. Relative accuracy was based on the mean percentage error (difference between actual and observed divided by actual multiplied by 100) and frequency of cases $\geq 5\%$, $\geq 10\%$, and $\geq 25\%$ error per device. Absolute precision was based on the standard deviation for the mean of actual steps recorded per device, whereas relative precision involved the coefficient of variation per device. We examined Spearman rho rank-order correlations between body mass index (i.e. crude measure of body fatness based on mass (kg) divided by height (m) squared), usual walking speed, and disability status (i.e. MSWS-12 and EDSS scores) with estimates of relative accuracy for examining possible sources of device/application inaccuracy. The level of significance for interpreting correlations as significant was set at 0.001 based on a Bonferroni correction. The analyses were performed in IBM SPSS Statistics for Windows (version 22; IBM SPSS Inc., Armonk, NY, USA).

Results

Table 3 presents the data for the relative and absolute metrics of accuracy and precision per device and application in the first trial of walking 500 steps;

Table 4 presents the same data for the second trial of walking 500 steps. Figure 1 illustrates the absolute accuracy and precision of the sensors and applications for the first 500-step trial, and Figure 2 provides the absolute accuracy and precision for the second 500-step trial. Regarding wearable sensors, the Jawbone UP2 and the Fitbit One had the best absolute and relative accuracy, and the Fitbit One had the best absolute and relative precision. The Fitbit One further had the fewest cases of $\geq 5\%$ error. Regarding applications, the Health and Health Mate had the best absolute and relative accuracy, and the Health application had the best absolute and relative precision. This pattern was repeatable between the two trials of walking 500 steps. There were statistically significant relationships between usual walking speed and the relative accuracy of the Yamax Digi-Walker for the first trial ($r_s = -.58$, $p < .001$) and the second trial ($r_s = -.63$, $p < .001$), and the Fitbit One for the first trial ($r_s = -.53$, $p < .001$) and the second trial ($r_s = -.53$, $p < .001$), as reported in Tables 5 and 6.

Discussion

This study provides a novel investigation of the accuracy and precision of commercially available motion sensors and smartphone applications in persons with mild and moderate MS, and extends previous research performed on healthy adults into a population with a neurological condition influencing ambulation.[3] Overall, the Fitbit One had the best absolute and relative accuracy and precision of all eight applications and devices across repeated trials, and the results regarding accuracy are entirely consistent with previous research involving healthy adults.[3] We extend that research by demonstrating that the Fitbit One had the best relative and absolute precision. Such evidence is important, as the Fitbit

Table 3. Accuracy and precision of sensors and applications in the first trial of walking 500 steps.

	Accuracy						Precision	
	Mean	95% CI	% error	$N \geq 5\%$ error	$N \geq 10\%$ error	$N \geq 25\%$ error	Standard deviation	Coefficient of variation
Digi-Walker SW-200 pedometer (Yamax)	457.3	428.5–486.0	8.5	11	9	5	95.7	0.2
UP2 (Jawbone)	480.5	467.3–493.7	3.9	18	10	2	43.9	0.1
UP Move (Jawbone)	457.8	427.9–487.8	8.4	21	12	3	99.6	0.2
Flex (Fitbit)	431.2	408.5–453.8	13.8	22	19	13	75.4	0.2
One (Fitbit)	490.6	485.6–495.5	1.9	4	2	0	16.4	0.0
Health (Apple)	486.4	474.2–498.6	2.7	13	3	2	40.7	0.1
Health Mate (Withings)	482.4	458.5–506.3	3.5	11	3	2	79.7	0.2
Moves (ProtoGeo Oy)	429.0	393.4–464.5	14.2	27	15	7	118.4	0.3

Table 4. Accuracy and precision of sensors and applications in the second trial of walking 500 steps.

	Accuracy						Precision	
	Mean	95% CI	% error	$N \geq 5\%$ error	$N \geq 10\%$ error	$N \geq 25\%$ error	Standard deviation	Coefficient of variation
Digi-Walker SW-200 pedometer (Yamax)	451.4	419.3–483.5	9.7	13	8	5	106.9	0.2
UP2 (Jawbone)	490.5	457.2–523.8	1.9	15	8	2	110.9	0.2
UP Move (Jawbone)	455.5	432.4–478.5	8.9	21	12	2	76.8	0.2
Flex (Fitbit)	438.1	418.0–458.2	12.4	24	20	8	67.0	0.2
One (Fitbit)	492.2	487.6–496.8	1.9	4	2	0	15.4	0.0
Health (Apple)	485.3	470.2–500.4	2.9	20	7	1	50.3	0.1
Health Mate (Withings)	492.4	474.5–510.2	1.5	21	9	1	59.5	0.1
Moves (ProtoGeo Oy)	437.6	395.1–480.1	12.5	29	20	7	141.5	0.3

Error bars indicate SD
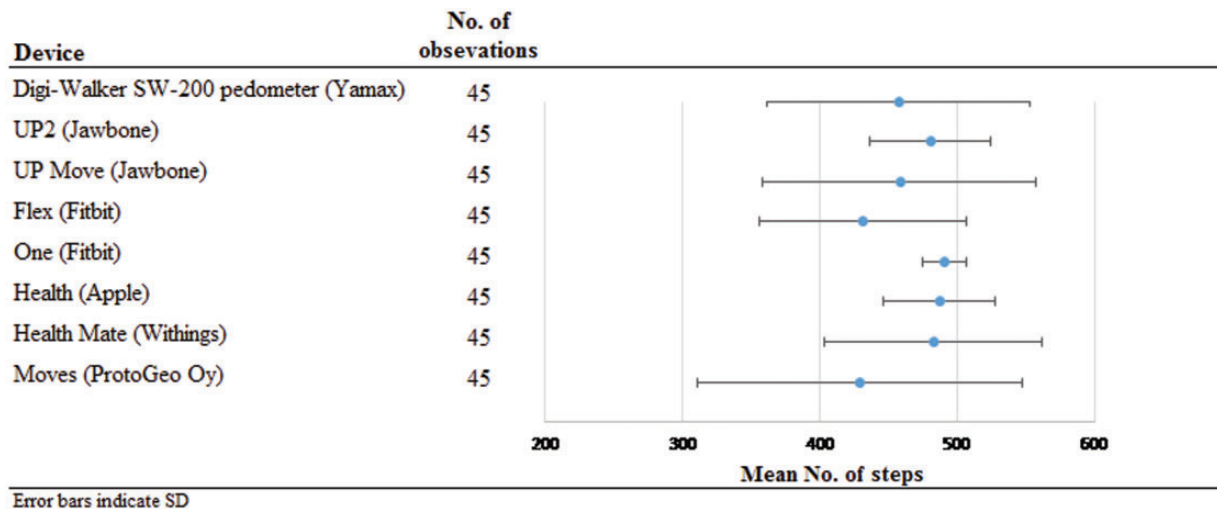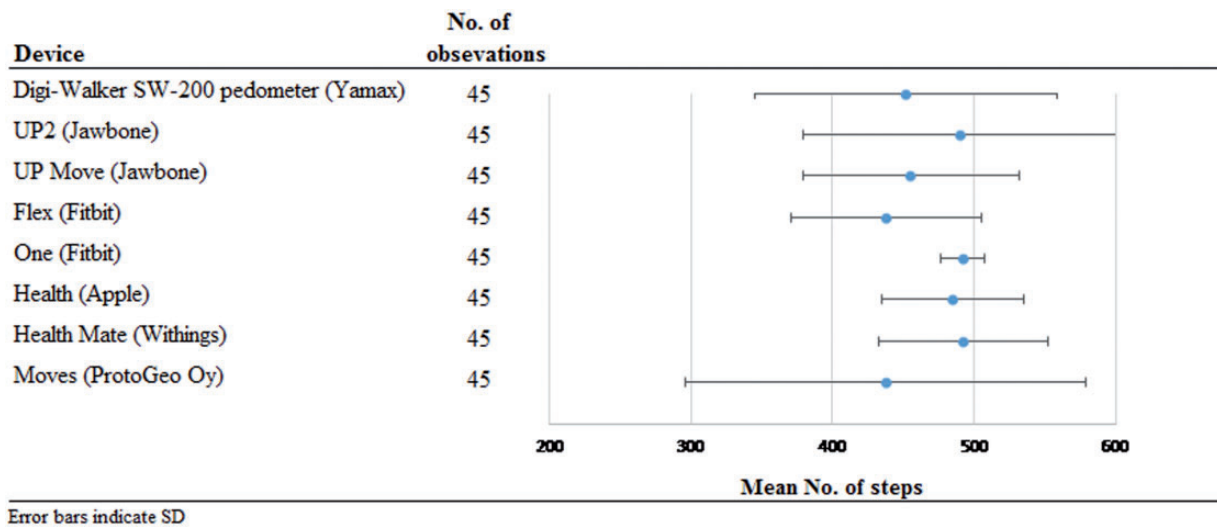
Figure 1. Absolute accuracy and precision of the sensors and application for measuring 500 steps in the first trial.

Error bars indicate SD

Figure 2. Absolute accuracy and precision of the sensors and application for measuring 500 steps in the second trial.

Table 5. Spearman Rho correlations of BMI, walking speed, and disability and relative accuracy of devices in the first trial of walking 500 steps.

	BMI	Usual walking speed	MSWS-12	EDSS
Digi-Walker SW-200 pedometer (Yamax)	0.28	−0.58*	0.23	0.21
UP2 (Jawbone)	0.06	−0.13	−0.16	0.02
UP Move (Jawbone)	0.18	−0.18	0.03	−0.05
Flex (Fitbit)	−0.15	−0.09	−0.06	−0.02
One (Fitbit)	−0.22	−0.53*	0.16	0.22
Health (Apple)	0.31	−0.11	−0.05	−0.17
Health Mate (Withings)	0.28	−0.23	−0.02	0.04
Moves (ProtoGeo Oy)	−0.31	−0.12	0.13	0.24

*indicates correlation is significant at 0.001 level

Table 6. Spearman Rho correlations of BMI, walking speed, and disability and relative accuracy of devices in the first trial of walking 500 steps.

	BMI	Usual walking speed	MSWS-12	EDSS
Digi-Walker SW-200 pedometer (Yamax)	−0.07	−0.63*	0.29	0.23
UP2 (Jawbone)	−0.06	−0.05	−0.01	0.06
UP Move (Jawbone)	−0.13	0.07	−0.01	0.12
Flex (Fitbit)	−0.17	−0.14	0.07	0.06
One (Fitbit)	−0.22	−0.53*	0.16	0.22
Health (Apple)	0.08	0.00	−0.11	−0.39
Health Mate (Withings)	−0.09	0.10	−0.20	−0.40
Moves (ProtoGeo Oy)	−0.27	−0.34	0.16	0.20

*indicates correlation is significant at 0.001 level

One has the necessary accuracy and precision for measurement of steps taken in clinical research and practice involving persons with MS. This complements the recent application of the Fitbit One by Biogen using PatientsLikeMe, wherein the device was feasible for measuring community-based, ambulatory physical activity in MS.[2] That study further provided data that patients considered the Fitbit One as important and practical for self-managing one's disease, and collectively the data on feasibility, accuracy, and precision have obvious relevance for integration of the Fitbit One into clinical research and practice.

We noticed variation in metrics of accuracy and precision across the motion sensors and applications. Some of this variation might be based on the bodily location of the devices. For example, some devices (e.g. Digi-Walker SW-200 (Yamax), UP Move (Jawbone), and One (Fitbit)) are worn on the waist and centered on the participant's non-dominant hip, and this is an ideal location for capturing

ambulatory physical activity that involves displacement of the center of mass. Other devices are worn on the wrist (UP 2 (Jawbone) and Flex (Fitbit)), and this is not an ideal location for capturing ambulatory physical activity that involves displacement of the center of mass. Indeed, a study examining step outputs obtained from waist and wrist accelerometer attachment sites while running/walking on a treadmill in a healthy population reported that wrist-worn sensors consistently detected fewer visually counted steps than the wrist attachments at most speeds, regardless of applied algorithm.[6]

Another possible influence on accuracy and precision is the technology of the wearable sensors. For example, both the Fitbit and Jawbone sensors use triaxial accelerometers, whereas the Yamax Digi-Walker pedometer uses a spring-loaded lever arm. The technology of the Yamax Digi-Walker makes it one of the least expensive sensors on the market, but its accuracy and precision are not as strong as those of the UP2 (Jawbone), UP Move (Jawbone),

and One (Fitbit) triaxial sensors. There was some variation across the smartphone applications. All smartphones utilize the M7 motion coprocessor of the iPhone, and this suggests that a difference in proprietor algorithms contributes to the varying differences in accuracy and precision between applications.

There were significant and strong associations between walking speed and the accuracy of the Digi-Walker (Yamax) and the One (Fitbit), whereby slower walking speed was associated with larger relative error in these two waist-worn sensors. Of note, the average speed of walking trials that yielded cases with ≥10% error for the Digi-Walker and Fitbit One was 1.9 mph and 1.6 mph, respectively, as a post-hoc examination of the actual speed wherein there is a significant deterioration of accuracy. This is consistent with previous research on the accuracy of the Digi-Walker (Yamax) in persons with MS, wherein larger inaccuracy occurred below 2.0 mph.[7] These data suggest that researchers should use caution when applying these devices to particularly slowly walking persons with MS, as such persons might have significant inaccuracy in the device output due to altered gait and greater perceived impact of walking impairment. To that end, future research should discern the exact point of departure for the accuracy of these devices through an evaluation of the devices against actual steps taken in a laboratory setting under a range of controlled manipulations of walking speed.

The observed decline in device accuracy across slower speeds is consistent with the body of research on the use of accelerometers in MS. One study reported minimal inaccuracy (4.1% error rate) while walking on a treadmill at 2.0 mph for the ActiGraph model 7164 accelerometer worn around the waist in persons with mild MS.[8] Another study examined the accuracy of the ActiGraph model GT3X+ accelerometer and the StepWatch Activity Monitor in persons with MS who had varying levels of disability, and observed that in the slow walking condition (i.e. 0.5 mph slower than a participant's comfortable walking speed), the StepWatch measured a greater percentage of actual steps taken (95.7%) than the ActiGraph (87.3%) in those with severe disability.[9] Thus, even research-grade devices have some problems with accuracy under slow walking conditions, particularly in those with the greatest burden of disease, and researchers should be aware of this when selecting a device for clinical research and practice.

This study is limited in that walking trials were conducted on a motorized treadmill and not overground. Importantly, we included an accommodation period of walking and established the speed based on T25FW at a comfortable walking speed, but treadmill walking might still be novel and not mimic real-life walking conditions. Indeed, real-world walking can be characterized by variable distance, speed, steps, terrain, and course, and this was not mimicked on the treadmill. Accordingly, future research is needed to confirm the accuracy and precision of these sensors and applications in persons with MS during over-ground walking that mimics free-living conditions, but our data provide a first examination of accuracy under controlled conditions before moving into the real world. Further, the study only included individuals with a self-reported ability to walk 500 steps without assistance (i.e. mild and moderate disability) and mostly relapsing–remitting multiple sclerosis (RRMS). Previous research has indicated decreased accuracy of sensors in persons with higher levels of disability,[9] and as such, it is necessary to examine new sensors in persons with mild walking impairment prior to expanding the research to persons with more severe disability. These results may not be generalizable broadly among persons with MS who use an assistive device for ambulation and who have progressive clinical courses. This, too, should be a focus in future research.

Commercially available, wearable sensors and applications for tracking free-living, ambulatory physical activity are soaring in popularity. These sensors and applications vary in numerous characteristics, such as wear site and price, as described in Table 2. We believe it is important to consider the accuracy and precision of these sensors and applications when implemented in research and clinical practice involving persons with MS. Our results suggest that the waist-worn One (Fitbit) has the highest accuracy and precision of the evaluated sensors and applications. Researchers and clinicians should consider these results when choosing a commercially available motion sensor or application for use in persons with MS.

References

1. Motl RW, Learmonth YC, Pilutti LA, et al. Top 10 research questions related to physical activity and multiple sclerosis. *Res Q Exerc Sport* 2015; 86: 117–129.

2. McIninch J, Datta S, DasMahapatra P, et al. Remote tracking of walking activity in MS patients in a real-world setting (P3.209). *Neurology* 2015; 84(14 Supplement): P3.209.

3. Case MA, Burwick HA, Volpp KG, et al. Accuracy of smartphone applications and wearable devices for tracking physical activity data. *JAMA* 2015; 313: 625–626.

4. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: An Expanded Disability Status Scale (EDSS). *Neurology* 1983; 33: 1444–1452.

5. Motl RW and Snook EM. Confirmation and extension of the validity of the Multiple Sclerosis Walking Scale-12 (MSWS-12). *J Neurol Sci* 2008; 268: 69–73.

6. Tudor-Locke C, Barreira TV and Schuna JM. Comparison of step outputs for waist and wrist accelerometer attachment sites. *Med Sci Sports Exerc* 2015; 47: 839–842.

7. Motl RW, McAuley E, Snook EM, et al. Accuracy of two electronic pedometers for measuring steps taken under controlled conditions among ambulatory individuals with multiple sclerosis. *Mult Scler* 2005; 11: 343–345.

8. Motl RW, Snook EM and Agiovlasitis S. Does an accelerometer accurately measure steps taken under controlled conditions in adults with mild multiple sclerosis? *Disabil Health J* 2011; 4: 52–57.

9. Sandroff BM, Motl RW, Pilutti LA, et al. Accuracy of StepWatch^TM and ActiGraph accelerometers for measuring steps taken among persons with multiple sclerosis. *PLoS ONE* 2014; 9: e93511.