International Journal of EXERCISE SCIENCE

Original Research

Wrist-worn Physical Activity Trackers Tend to Underestimate Steps During Walking

TIFFANY SEARS*[1], ELMER AVALOS*[1], SAMANTHA LAWSON*[1], IAN MCALISTER*[1], CHRIS ESCHBACH‡[2], and JENNIFER BUNN‡[1]

[1]Campbell University, Exercise Science, Buies Creek, North Carolina, USA; [2]Valencell Inc. Raleigh, North Carolina

*Denotes undergraduate student author, ‡Denotes professional author

ABSTRACT

International Journal of Exercise Science 10(5): 764-773, 2017. The purpose of this study was to determine step-count accuracy of pedometers at different walking speeds. Ten recreationally active participants walked at five treadmill speeds (0.89, 1.11, 1.34, 1.56, and 1.79 m/s) for five minutes while wearing four wrist-worn activity trackers (Fitbit Charge HR®, Garmin Vivosmart HR®, Apple iWatch®, Jawbone UP3®) and the hip-worn Digi-Walker®. Each step was manually counted by a research technician (benchmark). Total step count at each speed was obtained for each device and compared to the benchmark using one-way MANOVA and Pearson correlation coefficient. For all five speeds, the Digi-Walker® yielded the most accurate values, averaging -0.4% difference from the benchmark counted steps, and showed the strongest correlation, $r > .730$, $p < .05$, at every speed. The Fitbit averaged the highest percent difference of -10.2% from the benchmark of counted steps, and underestimated steps at all speeds ($p < 0.05$). Garmin averaged a -2.7% step difference, Jawbone averaged a -5.3% step difference, and the iWatch showed a -7.9% step difference. Specifically, the Fitbit, Garmin, and Jawbone got progressively worse with increasing speed, whereas the iWatch performed the worst at the slowest and fastest speeds. All wrist-worn devices tested tended to underestimate steps. These data indicate that wrist-worn pedometers are inaccurate even with a specific designed purpose: count steps in a controlled manner. Because these devices are inaccurate in this setting, they remain highly questionable for accuracy in a real-world setting in which the definition of a "step" becomes less finite.

KEY WORDS: Exercise movement techniques, telehealth standards, exercise tests

INTRODUCTION

It is estimated that for every 30 minutes of brisk walking, an individual takes between 3100 and 4000 steps, depending on the age and population demographics of the subject (6). Many governing bodies, like the American Heart Association and American College of Sports Medicine, suggest taking at least 10,000 steps a day for optimal health and to achieve minimum physical activity recommendations (15, 21, 28). Thus there is growing popularity in wearable physical activity trackers to allow the public to measure their movement. Pedometers and fitness trackers that work through accelerometry are both popular devices used to track daily step count. Since the creation of pedometers, they have been used by the public to count steps and have been utilized in research (23). Three major types of pedometers exist: the spring-suspended horizontal lever arm, the glass-enclosed magnetic reed proximity switch, and an accelerometer consisting of a horizontal beam and a piezoelectric crystal (6).

Previous studies have examined accuracy of pedometers with inconclusive results (22). Device accuracy varied due to factors including, but not limited to: the location of device on the body, the subjects body mass index (BMI), body fat percentage, the surface in which the individual is walking on, and if the subject utilizes a prosthetic or not (1, 3, 9, 10, 13, 16, 18, 28, 30). Although hip-worn devices have demonstrated better accuracy than devices in alternate positions it is important to note that placement alterations, such as using the posterior midline versus the anterior midline of the right thigh, can also affect accuracy of those devices at least in obese individuals (1). Of the studies that reveal pedometers are accurate at measuring step counts, most illustrate a decrease in accuracy as walking speed increases (7, 10, 19, 21, 26, 27, 30). Abnormal movements such as agility drills, skipping, hopping, and galloping have also been examined for step counting accuracy, and the results show a larger difference in calculated steps versus actual steps (24, 25).

Research indicates that the variance in step count accuracy is due to the reduced vertical acceleration of the hip that is not registered as a step in the device under test (6). These reductions in acceleration may be caused by a multitude of factors, including but not limited to, the use of prosthetics and/or the differences in gait activity. If an individual walks with impairment or a prosthetic of some sort, there may be a reduction in acceleration due to the lack of synced step acceleration where there may be less leg or arm swing and no activation of the device. Some devices, such as the Garmin and AcitGraph have incorporated the use of Global Positioning Systems (GPS) and accelerometers in an attempt to help correct any inaccuracies in these devices. However, these devices still underestimate actual steps taken at speeds less than 0.9 m/s (2.0 mph) (31). Fitbit is one of the industrial leaders in the development of new personal fitness device technology and is believed to be one of the most valid and reliable devices on the market (2, 7, 8, 11, 12, 14, 15, 29). Of the Fitbit pedometers, the most tested was the Fitbit One®, the only hip-based Fitbit (8). While testing the Fitbit One® and a wrist-based physical activity tracker Fitbit Flex®, the Fibit One® outperformed the Flex® during walking and running (8). This research further illustrates the increased accuracy at the hip versus alternate locations such as the wrist.

Despite a range of reported accuracies from commercially available pedometers consumers continue to shows a great deal of interest in them, particularly wrist-worn devices (8). Wrist-worn devices can be affected by additional variables that other types of pedometers may not, such as the location of the device placed on the wrist, how tightly the device is worn, and on which arm the device is placed (17). With more variables affecting wrist worn devices, more work must be done to ensure the best quality and accuracy. The objective of this study was to determine the step-count accuracy of commercially available wrist-worn pedometers (Fitbit Charge HR, Garmin VivoSmart, Apple iWatch, and Jawbone UP3) at different walking speeds. The authors hypothesize that the wrist-worn devices and hip device will be within ±3% error of the actual step count.

METHODS

Participants
Ten participants were recruited for the study who were between the ages of 18 and 40 years old, recreationally active, and low- or moderate-risk for cardiovascular disease according to the American College of Sports Medicine (ACSM) guidelines. Five males and five females completed the study (23.3 ± 5.23 years, 73.27 ± 10.94 kg, 75.26 ± 6.1 cm, 17.09 ± 6.9% body fat). Each participant signed an informed consent form that was approved by the university Institutional Review Board, completed a Physical Activity Readiness Questionnaire (PAR-Q), and an ACSM risk assessment.

Protocol
Resting heart rate and blood pressure were taken for each participant. Height, weight, and body composition was recorded for each participant using the InBody 570 bioelectrical analysis machine (Biospace Co. LTD, Korea).

Five different activity trackers were used to determine number of steps taken during five different walking speeds on a Trackmaster treadmill (Full Vision, Inc., Newton, Kansas). The treadmill was calibrated prior to data collection . The trackers included four wrist-worn devices: the Fitbit Charge HR (Fitbit Inc., San Francisco, California), Garmin Vivosmart HR (Garmin Ltd., Olathe, Kansas), Apple iWatch (Apple Inc., Mongolia, China), Jawbone UP3 (Jawbone, San Francisco, California); and one hip-worn device: the Digi-Walker (Yamax Health & Sports, Inc., San Antonio, Texas). A maximum of two activity trackers were worn on the left wrist of each participant per trial to ensure proper placement of the wrist devices. The activity trackers were placed on the wrist in randomized order with one of the two devices being more proximal to the other. The Digi-Walker was placed on the left hip parallel to the walking surface. It was located just above the anterior superior iliac spine on each participant to ensure secure placement on the hip. Participants were allotted time to get used to each speed prior to data collection. A research technician set each speed and let each participant walk as long as they needed in order to normalize gait for accurate step counts before the pedometers were placed on the wrist or any steps were counted. Each participant walked for five minutes at each of the five speeds: 0.89 (2.0 mph), 1.11 (2.5 mph), 1.34 (3.0 mph), 1.56 (3.5 mph), and 1.79 m/s (4.0 mph). Before the participant started walking for each speed, the speed was set by a research technician and the participant then stepped onto the treadmill to only have steps counted at accurate speeds. No speeds progressively increased or decreased as there was always a sudden stop and start by getting off and back onto the treadmill. Each participant did this twice during the same testing session in order to collect data on all five pedometers. A research technician manually counted steps during each trial using a handheld tally counter to serve as the benchmark. The Digi-Walker was reset after each walking speed was counted, but the wrist pedometers were not counting steps between speeds as the participants were standing still on the sides of the treadmill. The starting steps as well as the ending steps for each speed were documented for each of the wrist pedometers during the trials. A video recording of the participant's lower extremities only was also used to ensure benchmark accuracy. If the research technicians manual counting did not match the steps that were counted on the video recording, the video recording was used as it was proof of accuracy.

Statistical Analysis

Data were analyzed using the Pearson correlation coefficient to determine step count agreement between the benchmark and each activity tracker. Individual one-way multivariate analysis of variance (MANOVA) was also used to compare the benchmark to each device at the five different speeds. Percent differences between the benchmark and each activity tracker was also determined at each speed and overall. An alpha level of 0.05 was used to determine statistical significance. All data were analyzed using the SPSS version 19.0 (IBM, Chicago, Illinois).

RESULTS

Table 1 represents the results of the comparisons between each device and the benchmark. Wrist-worn devices tended to underestimate steps; exceptions were the Jawbone at 0.89 m/s and the Garmin at 1.11 m/s. The hip-worn Digi-Walker was the most accurate device tested, as it was the only tracker to have no significant differences from the benchmark (*Lambda*(5,14) = .668, p = .654). The MANOVA revealed a significant difference with the benchmark for the Fitbit Charge HR (*Lambda*(5,14) = 2.995, p = .048), the Garmin Vivosmart (*Lambda*(5,14) = 3.972, p = .019), the Jawbone UP3 (*Lambda*(5,14) = 4.953, p = .008), and the Apple iWatch (*Lambda*(5,14) = 3.572, p = .027). More specifically, the Fitbit was significantly different from the benchmark at each of the five walking speeds. The iWatch tested significantly different from the benchmark at the slowest and fastest walking speeds of 0.89 and 1.79 m/s. The Jawbone and Garmin devices tended to be more accurate at slower walking speeds.

Table 1. Mean ± SD step difference between each tracker and the benchmark. Negative values represent an underestimation in steps in the tracker compared to the benchmark. * indicates significant difference from the benchmark according to the paired sample t-test, p <.05.

Speed (m/s)	Fitbit	Jawbone	iWatch	Garmin	Digi-Walker
0.89	-37.8 ± 30.6* (p =.013)	.40 ± 46.9	-75.6 ± 96.2* (p =.013)	-6.4 ± 33.3	-13.7 ± 34.6
1.11	-28.9 ± 41.0* (p = .030)	-12.8 ± 37.2	-13.2 ± 82.9	6.3 ± 27.7	13.5 ± 26.8
1.34	-45.0 ± 46.5* (p =.040)	-20.8 ± 73.7	-14.0 ± 53.5	-9.8 ± 26.8	-14.4 ± 39.3
1.56	-78.4 ± 75.6* (p =.003)	-48.9 ± 88.6	-6.3 ± 57.0	-33.2 ± 42.8* (p =.037)	3.7 ± 17.7
1.79	-98.7 ± 87.7* (p =.003)	-72.7 ± 65.6* (p =.001)	-106.5 ± 73.1* (p < .001)	-36.4 ± 88.5	2.4 ± 10.8

The percent differences between each device and the benchmark are shown in Figure 1. Most devices became less accurate as the speeds increased, with the exception of the Digi-walker. The Digi-Walker yielded the most accurate, with a mean -.4% difference compared to the benchmark. Following the Digi-Walker, the Garmin showed to be the most accurate with a mean -2.7% difference. The Jawbone displayed a mean -5.3% difference while the iWatch displayed a mean -7.9% difference. The Fitbit showed to be the least accurate with a mean -10.2% difference of actual values.

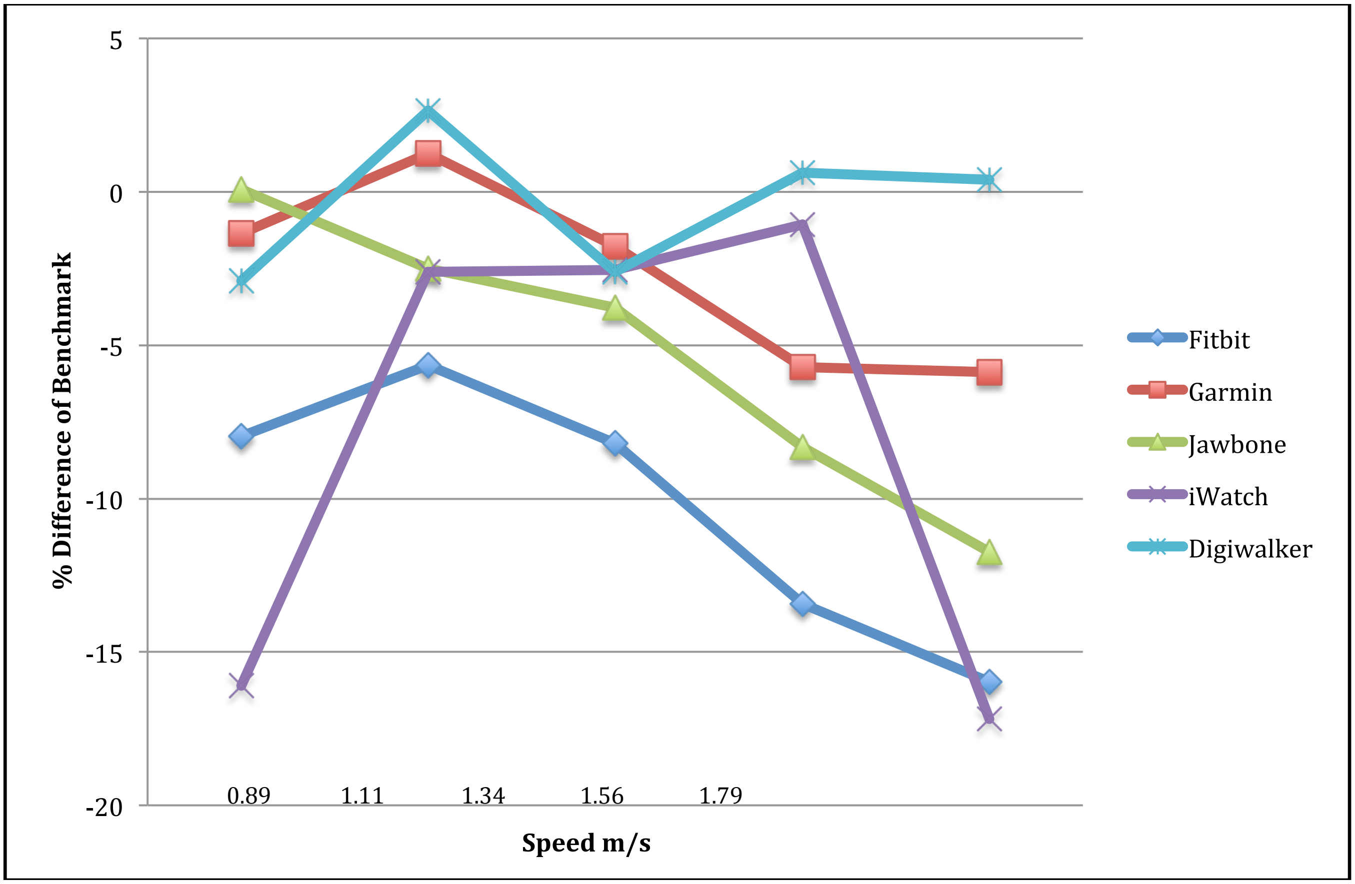

Figure 1. Percent difference of actual steps for the five speeds during treadmill walking.

Table 2 shows the results for the Pearson correlation. The Digiwalker showed the strongest correlation with high r-values at every speed along with all speeds showing statistically significant relationships ($p = 0.015$). The Fitbit displayed weak correlations with low r-values. The Garmin had moderate correlations and showed statistically significant relationships at 1.07 ($p = 0.046$) and 1.3 m/s ($p = 0.021$). Jawbone showed a weak correlation and had no statistically significant relationship ($p > 0.05$). Finally, the iWatch displayed weak correlations besides at 1.3 m/s ($p = 0.005$) that showed a high correlation statistically significant relationship.

Table 2. Pearson correlation r-values between each device and the benchmark for the five speeds walked on the treadmill. * denotes a significant relationship, p <.05.

Speed (m/s)	Fitbit	Garmin	Jawbone	iWatch	Digi-Walker
0.89	0.502	0.626	0.097	-0.285	0.828* (p =.003)
1.07	-0.114	0.640* (p =.046)	0.551	0.245	0.783* (p =.007)
1.30	0.520	0.713* (p =.021)	-0.272	0.810* (p =.005)	0.739* (p =.015)
1.52	-0.093	0.156	-0.313	0.302	0.901* (p <.001)
1.74	0.084	0.053	-0.301	0.056	0.964* (p <.001)

DISCUSSION

Accuracy of fitness trackers is important so as to avoid misleading the consumer. According to the Japanese industrial standards, the maximal permissible error of miscounted steps is 3% for pedometers to be allowed for commercial use (15). Based on this study, the most accurate pedometer tested was the Digi-Walker, a hip-worn device, which had less than -1% difference of actual steps from the benchmark. According to new United States technological standards, a device may have up to 10% error during walking, jogging, and running (20). All devices fell within these standards except for the Fitbit.

Despite the popularity of the Fitbit, our data shows the Fitbit Charge HR was the most error-prone device. In fact, it did not fall within the 3% difference reference used by the Japanese for any speed. The Fitbit underestimated actual step count at every speed tested with a mean of -10.2% difference, with lower error at slow speeds, and higher errors at faster speeds. This can cause some concern to the individuals whom like to do more running in their daily workout as it would increase the error at running speeds making their step count significantly lower than it should be. These data agree with a previous study indicating that the Fitbit has increasing error as gait speed increases (8). As an average, it takes an individual 2,000 steps to walk one mile, and if the individual were walking with the Fitbit that was tested, the step count would be 1,796, or approximately 204 steps off. The Garmin Vivofit as the most accurate of the wrist-worn devices for the speeds tested in this study, and falls within the Japanese industrial standards of 3% error miscounting of steps. However, when looking specifically at the speeds, the Garmin only falls within this standard for the three slowest speeds. Reasons for this increasing errors is unknown since algorithms of the Garmin device, and all other devices are unknown to the researchers. The iWatch tested well at all speeds except the slowest and the fastest speed of 0.89 and 1.79 m/s, with an average of -7.9% difference in step counts compared to those of benchmark. This could be due to the lack of research and need for iWatch to function as an activity tracker for its consumers because of its other and more popular features.

The most accurate speeds at which all devices performed at were between 1.11 and 1.34 m/s, and the accuracy of each device decreased at 1.79 m/s and higher speeds. One study displayed that the Digi-walker was found to have 97.8% accuracy on treadmill speeds of 1.34 m/s and 1.56 m/s, which can correlate to this study determining the results found were similar to others (14). This may be due to different technology used in each device while it has been suggested that this error in pedometers at slow walking speeds may come from reduced vertical acceleration at the hip (5). For pedometers that are worn on the wrist, the inaccuracy at the slower speeds may be due to the decreased arm swing. The data collected from the Digi-Walker follows this trend of decreased accuracy at slower walking speeds, indicating this hypothesis may stand true for most hip worn devices. However, wrist-based devices are not comparable to the same movement and force patterns of the hip, indicating there must be another cause of the inaccuracy seen at the wrist. Some wrist-based devices are meant to be worn tightly so that no excess movement may occur, while others are worn further up on the forearm to decrease hand and wrist movement. Each wrist-worn device in this study was placed on the wrist in accordance to the instructions provided by each company.

Algorithmic differences may also play a large factor in the accuracy of each device. Since each device uses a different algorithm, there may be a number of possible reasons why each device has different accuracies. With so many internal and external factors affecting device accuracy, it is hard to pinpoint one reason as to why each device measures step counts varies. Research using the Digi-Walker has shown that the mechanical pedometers detect foot strike induced acceleration as being a hip-worn device by using many spring-levered instruments within the device (13). Using hip-worn devices allow for vertical acceleration responses on the body needed to trigger an event in the device that the wrist-worn devices do not allow (5). Since wrist-worn devices do not get direct vertical acceleration from the foot strike, their accuracies vary greater than hip-worn devices as they need more advanced equipment to measure foot strikes. A study showed that the Digi-walker underestimated steps at slower speeds due to reduced ground impact forces at slower walking speeds (1).

Activity trackers can be used for feedback from physical activity to scientific research for health-based studies. While the underestimation of steps is not ideal for research purposes, it may be useful to individuals seeking to lose weight and live a healthier lifestyle. It is likely that these devices measure general motion quite well and may accurately measure metabolic equivalent tasks (METs), rather than specifically steps. METs could be useful and comparable to information provided on various pieces of exercise equipment, and as exercise prescription and exercise participation research is more frequently presented and reported in METs and MET-min. It is also likely that step count has been utilized because it is easily consumed by the lay public, in spite of the fact that the most common and popular devices measure steps rather poorly. Perhaps it would improve the information to the consumer if the devices provided these MET values in their application to provide accurate data, as well as a clinically relevant value to the user.

In conclusion, the wrist-worn physical activity monitors tended to underestimate steps, with the FitBit Charge HR performing the worst and the Garmin Vivosmart performing the best of the devices tested. With the release of the new technological standards for physical activity monitors (15), the Garmin Vivosmart was the only device tested to fall within these standards for every speed. While these data were collected in 2016, almost all of the companies that make these devices have released new or upgraded versions of physical activity monitors, and it would be important to examine the accuracy of these new devices as well because new technology and adjustments to algorithms have likely been made. Activity monitors are an important part of assessment and influence on behavior, and device accuracy should continue to be assessed.

REFERENCES

1. Abel MG, Hannon JC, Eisenman PA, Ransdell LB, Pett M, Williams DP. Waist circumference, pedometer placement, and step-counting accuracy in youth. Res Q Exerc Sport 80(3): 434-444, 2009.

2. Adam NJ, Spierer DK, Gu J, Bronner S. Comparison of steps and energy expenditure assessment in adults of fitbit tracker and ultra to the actical and indirect calorimetry. J Med Eng Technol 37(7): 456-462, 2013.

3. Ayabe M, Aoki J, Ishii K, Takayama K, Tanaka H. Pedometer accuracy during stair climbing and bench stepping exercises. J Sports Sci Med 7(2): 249-254, 2008.

4. Bassett DR, Wyatt HR, Thompson H, Peters JC, Hill JO. Pedometer-measured physical activity and health behaviors in united states adults. Med Sci Sports Exerc 42(10): 1819-1825, 2011.

5. Bassett DR, Leggett SR, Mathien CA, Main JA, Hunter DC, Duncan GE, et al. Accuracy of five electronic pedometers for measuring distance walked 310. Med Sci Sports Exerc 28(Supplement): 52, 1996.

6. Crouter SE, Schneider PL, Karabulut M, Bassett JDR. Validity of 10 electronic pedometers for measuring steps, distance, and energy cost. Med Sci Sports Exerc 35(8): 1455-1460, 2003.

7. De Cocker KA, De Meyer J, De Bourdeaudhuij IM, Cardon GM. Non-traditional wearing positions of pedometers: Validity and reliability of the omron HJ-203-ED pedometer under controlled and free-living conditions. J Sci Med Sport/Sports Med Australia 15(5): 418-424, 2012.

8. Diaz KM, Krupka DJ, Chang MJ, Peacock J, Ma Y, Goldsmith J, et al. Fitbit®: An accurate and reliable device for wireless physical activity tracking. Int J Cardiology 185: 138-140, 2015.

9. Duncan M, Black A, Yewdale J, Schoefield G, Mummery K. Accuracy of two makes of electronic pedometers for measuring step counts at various speeds. J Sci Med Sport 5(4): 23-23, 2002.

10. Feito Y., Bassett DR, Thompson DL, Tyo BM. Effects of body mass index on step count accuracy of physical activity monitors. J Phys Act Health 9(4): 594-600, 2012.

11. Ferguson T, Rowlands AV, Olds T, Maher C. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: A cross-sectional study. Int J Behav Nutr Phys Act 12(1): 42, 2015.

12. Fulk GD, Combs SA, Danks KA, Nirider CD, Raja B, Reisman DS. Accuracy of 2 activity monitors in detecting steps in people with stroke and traumatic brain injury. Phys Ther 94(2): 222, 2014.

13. Horvath S, Taylor DG, Marsh JP, Kriellaars DJ. The effect of pedometer position and normal gait asymmetry on step count accuracy. Appl Physiol Nutr Metab 32(3): 409-415, 2007.

14. Jung-Ming LEE, Youngwon KIM, Welk GJ. Validity of consumer-based physical activity monitors. Med Sci Sports Exerc 46(9): 1840-1848, 2014.

15. Kooiman TJM, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers. BMC Sports Sci Med Rehabil 7: 1-11, 2015.

16. Leicht AS, Crowther RG. Pedometer accuracy during walking over different surfaces. Med Sci Sports Exerc 39(10): 1847, 2007.

17. Mannini A, Intille SS, Rosenberger M, Sabatini AM, Haskell W. Activity recognition using a single accelerometer placed at the wrist or ankle. Med Sci Sports Exerc 45(11): 2193-2203, 2013.

18. Martin JB, Krc KM, Mitchell EA, Eng JJ, Noble JW. Pedometer accuracy in slow-walking older adults. Int J Ther Rehabil 19(7): 387-393, 2012.

19. Motl RW, McAuley E, Snook EM, Scott JA. Accuracy of two electronic pedometers for measuring steps taken under controlled conditions among ambulatory individuals with multiple sclerosis. Mult Scler 11(3): 343-345, 2005.

20. "New CTS Standards for IoT Advance Functionality of Tech Monitoring of Consumer Health." Accessed December 1, 2016. https://www.cta.tech/News/Press-Releases/2016/October/New-CTA-Standards-for-IoT-Advance-Functionality-of.aspx.

21. Nielson R, Vehrs PR, Fellingham GW, Hager R, Prusak KA. Step counts and energy expenditure as estimated by pedometry during treadmill walking at different stride frequencies. J Phys Act Health 8(7): 1004, 2011.

22. Schneider PL, Crouter SE, Lukajic O, Bassett DR. Accuracy and reliability of 10 pedometers for measuring steps over a 400-m walk. Med Sci Sports Exerc 35(10): 1779-1784, 2003.

23. Sigmundová D, Vašícková J, Stelzer J, Repka E. The influence of monitoring interval on data measurement: An analysis of step counts of university students. Int J Environ Res Public Health 10(2): 515-527, 2013.

24. Smith JD, Schroeder CA. Pedometer accuracy in elementary school children while walking, skipping, galloping, and sliding. Meas Phys Educ Exerc Sci 14(2): 92-103, 2010.

25. Stackpool CM, Porcari JP, Mikat RP, Gillette C, Foster C. The accuracy of various activity trackers in estimating steps taken and energy expenditure. J Fitness Res 3(3): 32-48, 2014.

26. Storm FA, Heller BW, Mazzà C. Step detection and activity recognition accuracy of seven physical activity monitors. PloS One 10(3): e0118723, 2015.

27. Swartz AM, Bassett DR, Moore JB, Thompson DL, Strath SJ. Accuracy of an electronic pedometer in adults with varying body mass index levels. Med Sci Sports Exerc 35(Supplement 1): S283, 2003.

28. Swartz AM, Strath SJ, Miller NE, Grimm EK, Ewalt LA, Loy MS, et al. Validity of physical activity monitors in assessing energy expenditure in normal, overweight, and obese adults. Open Sports Sci J 2(1): 58-64, 2009.

29. Takacs J, Pollock CL, Guenther JR, Bahar M, Napier C, Hunt MA. Validation of the fitbit one activity monitor device during treadmill walking. J Sci Med Sport 17(5): 496-500, 2014.

30. Webber SC, Magill SM, Schafer JL, Wilson KCS, Dept. of Physical Therapy, University of Manitoba, Winnipeg, MB, Canada. GT3X+ accelerometer, yamax pedometer and SC-StepMX pedometer step count accuracy in community-dwelling older adults. J Aging Phys Act 22(3): 334-341, 2014.

31. Webber SC, Porter MM. Monitoring mobility in older adults using global positioning system (GPS) watches and accelerometers: A feasibility study. J Aging Phys Act 17(4): 455, 2009.