2015

How Accurate Are Wearable Activity Trackers For Measuring Steps?

Jung-Min Lee
University of Nebraska at Omaha, jungminlee@unomaha.edu

Gregory Jones
University of Nebraska at Omaha

Hyun-Sung An
University of Nebraska at Omaha, han@unomaha.edu

Jungyoon Kim
University of Nebraska Medical Center

Danae M. Dinkel
University of Nebraska at Omaha, dmdinkel@unomaha.edu

UNIVERSITY OF NEBRASKA AT OMAHA

HOW ACCURATE ARE WEARABLE ACTIVITY TRAKCERS FOR MEASURING STEPS?

Jung-Min Lee[1]. Ph.D., Gregory Jones[1], Hyun-Sung An[1], Jungyoon Kim[2], Danae M. Dinkel[1]
[1]College of Health, Physical Education and Recreation, University of Nebraska at Omaha, [2]College of Public Health, University of Nebraska Medical Center

ABSTRACT

Wearable activity trackers have become popular for tracking individual's daily physical activity, but little or no information is available to substantiate the validity of these devices in step counts. **PURPOSE:** The purpose of this study was to systemically examine the validity of newly developed wearable activity trackers for measuring steps compared to the criterion measure (hand tally) in two different conditions. **METHODS:** Twenty (28.2±4.8 years) healthy males (n=19) and females (n=17) participated in the study. The participants were fitted with eight wearable activity trackers while walking and running on a treadmill (speeds of 2, 2.5, 3, 3.5, 4, and 5 mph) for 3-minutes at each speed. For overground protocol, participants walked at three-self-determined speeds; gradually becoming faster (slow, normal, and fast) for one lap on an indoor track (200 meter track). The number of actual steps taken was manually tallied by researchers using a hand-tally counter. The monitors included the Basis B1 band (BB), Misfit Shine (MS), Polar Loop (PL), and Jawbone UP (UP) worn on the right wrist; the Nike+Fuelband (NF), Garmin VivoFit (GV), and Fitbit Flex (FF) worn on the left wrist; and Withings Pulse (WP) and Fitbit Zip (FZ) worn with a clip on the waist. Step counts from each monitor were compared with criterion values from manually counted steps. **RESULTS:** Total step counts (means ± SD) were 329.5±71.0, 267.8±89.9, 290.6±105.1, 326.2 ±73.2, 282.2±85.1, 294.3±85.8, 329.2±70.0, 322.1±75.7, 310.8±82.8, and 318.±76.7, for manual counts, NF, MS, WP, PL, FF, FZ, UP, GV, and BB, respectively. Corresponding absolute error rates (computed as the average absolute value of the individuals' errors) were 19.8±16.4%, 18.9±12.2%, 17.4±15.8%, 11.3±13.1%, 0.7±1.4%, 4.5±7.8%, 6.6±12.6%, and 3.5±6.0%, respectively. ANOVA and Post hoc analyses with Bonferroni revealed the MS, WP, FZ, UP, GV, and BB were the devices to give non-significant differences (p> .05) compared to the manual step counts, but significant differences were found with NF, PL, and FF. **CONCLUSION:** The results demonstrate that the waist-oriented trackers, FZ and WP, show the most accuracy in measuring steps. However, promising preliminary findings were observed with the wrist-oriented trackers, BB, UP, and GV.

INTRODUCTION

- Wearable activity trackers have become popular for tracking individual's daily physical activity, but little or no information is available to substantiate the validity of these devices in step counts.

PURPOSE

- The purpose of this study was to systemically examine the validity of newly developed wearable activity trackers for measuring steps compared to the criterion measure (hand tally) in two different conditions.

METHODS

Participants

	Male	Range	Female	Range
Age (yrs)	35.4 ± 14.6	21.0 – 66.0	26.7 ± 5.9	19.0 – 39.0
Height (cm)	177.4 ± 5.9	163.8 – 186.7	168..3 ± 5.6	157.5 - 177.8
Weight (kg)	79.9 ± 12.1	57.0 – 103.9	63.4 ± 8.7	46.3 – 79.3
BMI (kg·m^{-2})	25.3 ± 2.9	20.3 – 30.9	22.3 ± 2.6	17.0 – 26.4

Abbreviations: BMI, Body mass index

- Twenty (28.2±4.8 years) healthy males (n=19) and females (n=17) participated in the study.

METHODS (Cont.)

Instruments

- The monitors included the Basis B1 band (BB), Misfit Shine (MS), Polar Loop (PL), and Jawbone UP (UP) worn on the right wrist; the Nike+Fuelband (NF), Garmin VivoFit (GV), and Fitbit Flex (FF) worn on the left wrist; and Withings Pulse (WP) and Fitbit Zip (FZ) worn with a clip on the waist.

Procedures

- The participants were fitted with eight wearable activity trackers while walking and running on a treadmill (speeds of 2, 2.5, 3, 3.5, 4, and 5 mph) for 3-minutes at each speed.
- For overground protocol, participants walked at three-self-determined speeds; gradually becoming faster (slow, normal, and fast) for one lap on an indoor track (200 meter track).
- The number of actual steps taken was manually tallied by researchers using a hand-tally counter.

Data Analyses

- The Mean step scores was calculated for each treadmill walking speed and overground walking speed for all eight pedometers.
- The scores were tested using one-sample t-tests and Bonferroni corrections for multiple comparisons.
- The Mean Absolute Percent Error (MAPE) were calculated to determine accuracy of each device.
- Correlations were be than calculated for each monitors steps recorded and the actual steps counted.
- Bland-Altman plots were used to distinguish similarities between the different measures of steps.

RESULTS

Table 2. Total Steps for treadmill and overground

Treadmill	N	Mean	Std. Deviation	Minimum	Maximum	MAPE (%)
Observed Steps	35	359.27	58.21	233.0	506.0	
Fitbit Zip	35	358.51	57.03	231.0	499.0	0.6
Withing Pulse	35	358.15	57.81	244.0	501.0	1.2
Jawbone UP	35	353.01	64.42	145.0	506.0	4.0
Basis B1	35	341.18	72.14	83.0	505.0	6.0
Garmin VivoFit	35	339.00	69.42	64.0	501.0	5.9
SenseWear Mini	35	337.81	74.99	97.9	499.0	6.9
Fitbit Flex	35	331.63	68.25	124.0	504.0	7.6
Misfit Shine	20	327.44	79.42	149.0	524.0	10.7
Polar Loop	35	314.17	81.40	68.0	538.0	16.1
Nike Fuel	35	286.06	98.93	31.0	494.0	23.3
Overground	N	Mean	Std. Deviation	Minimum	Maximum	MAPE (%)
Observed Steps	35	297.96	33.41	211.0	356.0	
Fitbit Zip	35	266.59	33.88	206.0	353.0	1.0
Withing Pulse	35	263.92	33.26	176.0	352.0	2.1
Jawbone UP	35	265.76	36.61	198.0	377.0	4.8
Basis B1	35	253.06	48.94	62.0	361.0	6.9
Garmin VivoFit	35	250.68	52.59	113.0	363.0	7.5
SenseWear Mini	35	245.86	34.65	153.0	358.0	8.0
Fitbit Flex	35	230.74	64.61	116.0	366.0	16.3
Misfit Shine	20	232.27	40.22	162.0	310.0	13.4
Polar Loop	35	231.45	50.73	116.0	350.0	17.7
Nike Fuel	35	202.93	48.99	21.0	297.0	22.5

* Added measured resting energy expenditure.

Table 3. Correlation matrix for total number of steps

Treadmill	Observed Steps	Fitbit Zip	Withing Pulse	Jawbone UP	Basis B1	Garmin VivoFit	SenseWear Mini	Fitbit Flex	Misfit Shine	Polar Loop	Nike Fuel
Observed Steps	1.00	.997**	.998**	.942**	.823**	.736**	.885**	.782**	.832**	.666**	.786**
Fitbit Zip		1.00	.985**	.939**	.816**	.726**	.883**	.785**	.815**	.662**	.797**
Withing Pulse			1.00	.936**	.825**	.719**	.877**	.777**	.848**	.659**	.787**
Jawbone UP				1.00	.798**	.706**	.872**	.804**	.801**	.702**	.782**
Basis B1 Band					1.00	.678**	.771**	.728**	.768**	.530**	.737**
Garmin VivoFit						1.00	.813**	.788**	.737**	.508**	.643**
Sesewear Mini							1.00	.714**	.742**	.677**	.778**
Fitbit Flex								1.00	.784**	.564**	.672**
Misfit Shine									1.00	.522**	.704**
Polar Loop										1.00	.586**
Nike Fuel											1.00

Overground	Observed Steps	Fitbit Zip	Withing Pulse	Jawbone UP	Basis B1	Garmin VivoFit	SenseWear Mini	Fitbit Flex	Misfit Shine	Polar Loop	Nike Fuel
Observed Steps	1	.987**	.933**	.824**	.633**	.801**	.496**	.753**	.423*	.502**	.297*
Fitbit Zip		1	.936**	.786**	.624**	.799**	.479**	.763**	.351	.506**	.244*
Withing Pulse			1	.802**	.594**	.758**	.454**	.717**	.394*	.507**	.338**
Jawbone UP				1	.383*	.663**	.496**	.612**	.224	.495**	.245*
Basis B1					1	.519**	.436**	.456**	.149	.244*	.258*
Garmin VivoFit						1	.445**	.808**	.440**	.492**	.301**
Sensewear Mini							1	.476**	.262	.369**	.453**
Fitbit Flex								1	.411**	.462**	.233*
Misfit Shine									1	.080	.238
Polar Loop										1	.196
Nike Fuel											1

** Correlation is significant at the 0.01 level (2-tailed) and * Correlation is significant at the 0.05 level (2-tailed).

RESULTS (Cont.)

Figure 1. Mean absolute percentage error (± SD) for all monitors with total steps (n = 38).

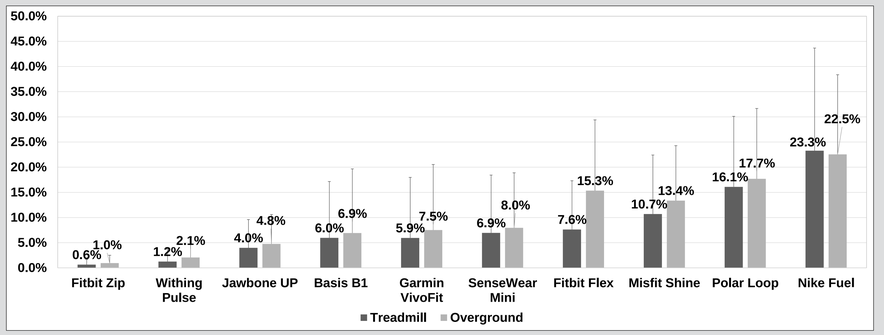
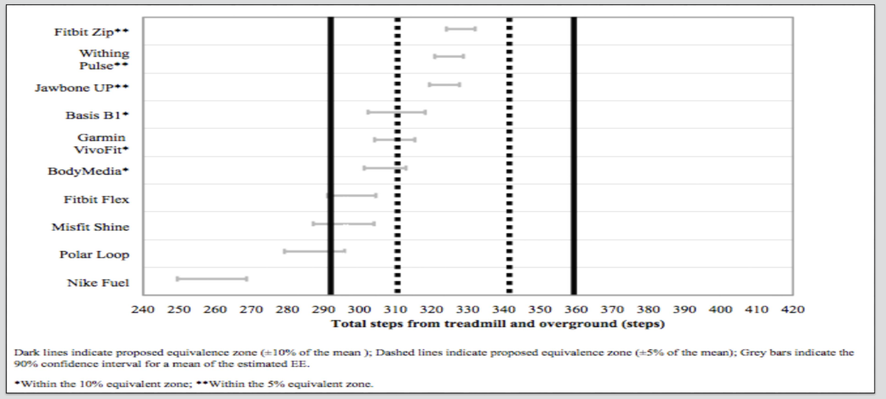


Figure 2. Results from 95% equivalence testing for agreement in total steps between observed steps and all monitors.



Dark lines indicate proposed equivalence zone (±10% of the mean); Dashed lines indicate proposed equivalence zone (±5% of the mean); Grey bars indicate the 90% confidence interval for a mean of the estimated EE.

*Within the 10% equivalent zone; **Within the 5% equivalent zone.

- ANOVA and Post hoc analyses with Bonferroni revealed the MS, WP, FZ, UP, GV, and BB were the devices to give non-significant differences (p> .05) compared to the manual step counts, but significant differences were found with NF, PL, MS, and FF.
- Fitbit zip, Withing Pulse, Jawbone UP, Basis B1, Garmin VivoFit, and BodyMedia Fit are in the 10% equivalent zone.
- Fitbit zip, Withing Pulse, Jawbone up are in the 5% equivalent zone.
- Jawbone UP is the only wrist-oriented monitor that is in the 5% equivalent zone.

DISCUSSION

- The present study demonstrated the waist-oriented trackers, Fitbit Zip and Withing Pulse, show the most accuracy in measuring steps.
- However, promising preliminary findings were observed with the wrist-oriented trackers, BB, UP, and GV.
- Additional research is needed to examine these trackers in free-living settings.