Validation of Utilizing a Modern Pedometer as a Measure of Patient Activity

Russell Stitzlein, MD
Alexander Neuwirth, MD
Keith Baldwin, MD MPH
Neil Sheth, MD

Introduction

Pedometers have been used for several years to measure the number of steps taken by users. Their recent popularization and advanced technologies such as Fitbit™ as well as digitized recording and applications linked to smartphones provides an opportunity for widespread clinical use by orthopaedic surgeons to evaluate patients pre-operatively and post-operatively.

Pedometers have been used with mixed results as a measure of patient activity in patients with total joint arthroplasty, with most studies reporting accuracy rates within 3%-5% of the true value.[1-7] Previous studies have found that following total joint arthroplasty, patients average approximately 5,000-7000 steps per day.[3,4] Age <60 years and male gender were found to be associated with increased steps per day. Multiple authors have commented on the wide variation in steps per day by patients undergoing total joint arthroplasty, with one group reporting a standard deviation of 3,040 steps per day[3] and another reporting a 15-fold difference in average daily steps among patients with the same UCLA activity scale score.[4]

Recent studies have been conducted to assess the use of smartphone pedometers in measuring patient activity. Results have been mixed, with most authors finding that smartphone application-based pedometers are too inaccurate for clinical evaluation/application.[8-10] Nolan, *et al*; however, showed that newer iPhone® models are 99% accurate in determining walking versus running and are able to accurately predict speed.[11] The iPhone® Health App calculates step-count and distance by integrating data obtained from the accelerometers, gyroscopes and compasses included in the phone. To date, no studies have reported on the accuracy or outcomes utilizing the iPhone® Health App available on iOS8 or later for the iPhone 5S or later smartphone models.

Hypothesis: Step counts as recorded by a modern pedometer (Health App on iPhone® 5S or later version) will correlate with the UCLA activity scale, Short Form (SF)-36 physical domain scores, Timed Up and Go (TUG) test scores, as well as Hip disability and Osteoarthritis Outcome Score (HOOS) and Knee injury and Osteoarthritis Outcome Score (KOOS) scores.

Specific Aims

1. Evaluate the accuracy of a modern pedometer (Health App).
2. Establish baseline values (T-score) for each of the outcomes measures and baseline pedometer step and distance values in a young healthy population.
3. Establish baseline values (Z-score) for each of the outcomes measures and baseline pedometer step and distance values in an elderly healthy population.
4. Correlate outcomes measures and pedometer step and distance values in patients with hip or knee osteoarthritis.

Methods

Institutional IRB approval was obtained prior to enrollment of subjects. A total of 110 subjects were enrolled in the preliminary study groups (validation and young healthy cohorts) with a plan to enroll 300 additional subjects for the final validation.

Patient Populations

1. Accuracy Cohort: Ten young healthy residents were recruited to assess the accuracy of a modern pedometer (Health App on iPhone® 5S or later version). Subjects all performed ten trials of 100 manually counted steps on both flat ground and on stairs. Pedometer data for each trial was then recorded for accuracy analysis.
2. Young Healthy Cohort: One hundred subjects aged 18-40 (medical students and residents) without chronic orthopaedic spine or lower extremity conditions were recruited to serve as young healthy control group. Each subject completed the UCLA activity scale, the SF-36, the HOOS and KOOS surveys and underwent a Timed Up and Go (TUG) test at the time of initial enrollment. The data from the Health App on the subject's smartphone was then extracted and the values for steps and distance for the 30-days immediately preceding the enrollment date were recorded and used for analysis.
3. Elderly Healthy Cohort *(Future Enrollment Group)*: This group will consist of 100 patients age ≥60-years old presenting to

orthopaedic surgery clinic with no history of chronic spine or lower extremity condition.

4. Hip Osteoarthritis Cohort *(Future Enrollment Group)*: This group will consist of 100 patients presenting to orthopaedic clinic with clinical and radiographic evidence of unilateral or bilateral **hip** degenerative joint disease with no ***clinical*** evidence of additional lower extremity joint or spine disease.
5. Knee Osteoarthritis Cohort *(Future Enrollment Group)*: This group will consist of 100 patients presenting to orthopaedic clinic with clinical and radiographic evidence of unilateral or bilateral **knee** degenerative joint disease with no ***clinical*** evidence of additional lower extremity joint or spine disease.

Results

Pedometer Accuracy

Five male and five female residents participated in the pedometer accuracy cohort. The male subjects had an average age of 29.4 years compared to 28.0 years for the female subjects ($p = 0.5$). Male subjects were an average of 72.2 inches tall compared to 63.8 inches for the female subjects ($p < 0.001$). Male subjects weighed an average of 203.0 pounds compared to an average of 121.4 pounds for female subjects ($p = 0.01$).

The accuracy analysis is summarized in Table 1. For all subjects, the pedometer recorded 2.1% more steps on flat ground than were manually counted by subjects. The discrepancy on flat ground for male subjects was +0.6% versus +3.5% for female subjects ($p = 0.26$). For all subjects, the pedometer recorded 5.9% more steps on stairs than were manually counted by subjects. The discrepancy on stairs for male subjects versus female subjects was statistically significant (1.6% vs. +10.2%, $p = 0.02$). Overall, there was no statistical difference between the pedometer data for all subjects on flat ground versus stairs ($p = 0.10$); this was also true for flat ground versus stairs for both the male ($p = 0.67$) and female ($p = 0.11$) subgroups.

Table 1. Pedometer accuracy meas ured as steps recorded on flat ground and on stairs when subjects manu ally counted 100 steps.

Flat Ground	N	Mean	STDEV	Range
Male	5	100.6	2.0	89–114
Female	5	103.5	5.0	90–128
Total	10	120.1	3.9	89–128
Stairs	**N**	**Mean**	**STDEV**	**Range**
Male	5	101.6	3 .0	86–105
Female	5	110.2	5.5	93–133
Total	10	105.9	6.1	86–133
Group	**vs**	**Group**	**p-value**	
Total Flat		Total Stain	0.10	
Male Flat		Male Stairs	0.67	
Female Flat		Female Stairs	0.11	
Male Flat		Female Flat	0.26	
Male Stairs		Female Stairs	0.02	

Young Healthy Cohort

Demographic analysis for the young healthy cohort is summarized in Table 2. Seventy-four males and twenty-six females participated. The cohort was predominantly Caucasian (87%). Twenty-four subjects reported a history of a prior lower extremity injury; no subjects reported an active acute or chronic lower extremity condition. Examples of prior injuries reported include meniscus tears, septic hip arthritis, metatarsal stress fracture, femoral shaft stress fracture, hip avulsion injury, pediatric femur fracture, pediatric tibial fracture, Osgood-Schlatter and aneurysmal bone cyst. Seventeen subjects reported prior lower extremity surgery. Examples of prior surgical procedures include knee arthroscopy, irrigation & debridement of open fracture, curettage and bone grafting and total hip arthroplasty.

Results of outcome instruments and pedometer data for the young healthy cohort is summarized in Table 3. The mean UCLA activity scale score was 8.9 (range, 5-10). Mean scores on the SF-36 were >95 for physical domains. Mean and median scores were lower for the energy/fatigue domain (62.0 and 65.0, respectively), the emotional well-being domain (82.2 and 84.0, respectively), and the general health domain (80.8 and 85.0, respectively). Mean scores for the HOOS were ≥ 89.9 for all domains. Median scores for all HOOS domains were 100. Mean scores for the KOOS were ≥ 96.6 for all domains.

Table 2. Demographic variables for the young healthy cohort.

Demographics	N	Mean	STDEV	Range
Gender				
Male	74			
Female	26			
Age		28.6	2.9	24.9–36.6
Height (inches)		70.2	4.5	56–77
Weight (lbs.)		171.8	37.3	105–265
BMI		24.3	3.3	19.2–33.4
Ethnicity				
Cacuasian (non-Hispanic)	87			
Caucasian (Hispanic)	3			
African-Amercan/Black	0			
Asian/Pacificl Islander	10			
Native American	0			
Prior Lower Extremity Injury	24			
Prior Lower Extremity Surgery	17			

Table 3. Outcomes measures by instrument and pedometer data for the young healthy cohort.

Outcome Measures	Mean	Median	STDEV	Range
UCLA Activity Scale	8.9	10	1.6	5–10
SF-36				
Physical Functioning	99.5	100	1.6	95.0–100
Role Limitations Due to Physical Health	100	100	0	
Energy/Fatigue	62.0	65.0	18.9	10.0–100
Emotional Well-Being	82.2	84.0	10.2	48.0–100
Social Functioning	96.8	100	8.1	62.5–100
Pain	95.	100	6.3	67.2–100
General Health	80.8	88	14.3	45.0–100
HOOS				
Pain	99.4	100	2.1	90.0–100
Symtpoms	89.9	100	2.5	90.0–100
Activities of Daily Living (ALDs)	100	100	0.3	98.5–100
Sports/Recreation	89.9	100	3.1	87.5–100
Quality of Life (QOL)	98.4	100	5.7	75.0–100
KOOS				
Pain	97.9	100	3.8	86.1–100
Symtpoms	96.8	100	5.0	85.7–1100
Activities of Daily Living (ALDs)	99.4	100	1.6	94.1–100
Sports/Recreation	96.6	100	7.6	70.0–100
Quality of Life (QOL)	97.8	100	5.5	75.0–100
Timed Up and Go (seconds)	5.50	5.35	1.40	2.55–9.52
Pedometer Data				
Steps	8118.2	7667.4	2257.5	4522–14478
Distance (meters)	6105.9	5697.4	1859.6	3067–11313
Distance (miles)	3.8	3.5	1.2	1.9–7.0

Median scores for all KOOS domains were 100. The mean TUG test time was 5.50 seconds. Subjects walked an average of 8118.2 steps [er day and 3.8 miles per day as recorded by the iPhone® Health App.

Discussion

The iPhone® Health App pedometer appears to accurately record steps and could represent a reliable way to measure patient activity. In our cohort, the iPhone® pedometer overestimated by 2.1% on flat ground and by 5.9% on stairs. If a larger study were to confirm the accuracy of the iPhone® Health App pedometer, it would be contrary to the findings of recent studies that have shown poor accuracy for smartphone pedometers.[8,9] Improved accuracy of iPhone® Health App pedometer could be due to decreased reliance on GPS data and/or improved data collection capabilities and programming in newer smartphones.

The SF-36 physical domain scores were high, on average, for the young healthy cohort as expected, with all physical domains displaying a mean >95. In comparison, the emotional domains demonstrated lower scores, with means from 62.0 for energy/fatigue to 82.2 for emotional well-being. At first glance, it might be surprising to find that there were deficits in the emotional domains on the SF-36, but a closer look at the domains and the study population provide a clear reasoning for these findings. The emotional domains of the SF-36 inquire about sleep habits, energy and stress levels. Given that our cohort exclusively included medical students and residents whose demanding academic and professional responsibilities result in chronic fatigue and stress, the lower scores are not unexpected.

The pedometer data for the young healthy controls demonstrated an average of 8,118.2 steps per day with a range of 4,522—14,478 steps per day. When compared to published step data on patients following total joint arthroplasty, it is somewhat surprising that our young healthy controls do not display a higher level of physical activity as recorded by the iPhone® Health App. It is quite possible that our mean value under-represents the level of physical activity of our subjects and of the general population of young healthy individuals. Potential reasons for this measured discrepancy could be that our subjects are actually more active but fail to keep their smartphones on their person during periods of exercise. If this is the case, it could under-represent the true mean by a significant factor. An alternative explanation is that our study population of medical students and residents, while young and healthy, are less active than they otherwise would be due the nature of their school and professional responsibilities.

As a whole, the data obtained from the young healthy cohort should serve as an effective baseline for future analysis after the enrollment of the final cohorts. The young healthy cohort will serve to establish a T-score, much the same way is done for bone density evaluation and the diagnosis of osteopenia and osteoporosis. The elderly healthy cohort will serve to establish a Z-score for physical activity. We anticipate that the utilization of advanced pedometers will gain popularity in orthopaedics as a means to objectively and non-invasively follow the activity levels of patients and serve as a quick and reproducible proxy for patient-report outcomes.

Conclusions

The iPhone® Health App Pedometer records steps walked with acceptable accuracy. Young healthy controls have high levels of activity as measured by the UCLA activity scale, have excellent physical health as evidence by high scores on SF-36 physical domains and near perfect scores on the HOOS and KOOS tools. This group can be used to establish a baseline for comparison in a manner similar to the assignment of the T-score in bone density screening and diagnosis of osteopenia and osteoporosis.

References

1. **Fukutani N, Iijima H, Aoyama T, *et al.*** Knee pain during activities of daily living and its relationship with physical activity in patients with early and severe knee osteoarthritis. *Clin Rheumatol.* April 2016. doi:10.1007/s10067-016-3251-8.

2. **Arnold JB, Walters JL, Ferrar KE.** Does Physical Activity Increase After Total Hip or Knee Arthroplasty for Osteoarthritis? A Systematic Review. *J Orthop Sports Phys Ther.* April 2016:1-42. doi:10.2519/jospt.2016.6449.

3. **Schmalzried TP, Szuszczewicz ES, Northfield MR, *et al.*** Quantitative Assessment of Walking Activity after Total Hip or Knee Replacement*. *J Bone Jt Surg Am.* 1998;80(1):54-59.

4. **Zahiri CA, Schmalzried TP, Szuszczewicz ES, Amstutz HC.** Assessing activity in joint replacement patients. *J Arthroplasty.* 1998;13(8):890-895.

5. **Shepherd EF, Toloza E, McClung CD, Schmalzried TP.** Step activity monitor: Increased accuracy in quantifying ambulatory activity. *J Orthop Res.* 1999;17(5):703-708. doi:10.1002/jor.1100170512.

6. **Silva M, McClung CD, dela Rosa MA, Dorey FJ, Schmalzried TP.** Activity Sampling in the Assessment of Patients With Total Joint Arthroplasty. *J Arthroplasty.* 2005;20(4):487-491. doi:10.1016/j.arth.2004.08.013.

7. **Silva M, Shepherd EF, Jackson WO, Dorey FJ, Schmalzried TP.** Average patient walking activity approaches 2 million cycles per year. *J Arthroplasty.* 2002;17(6):693-697. doi:10.1054/arth.2002.32699.

8. **Leong JY, Wong JE.** Accuracy of three Android-based pedometer applications in laboratory and free-living settings. *J Sports Sci.* March 2016:1-8. doi:10.1080/02640414.2016.1154592.

9. **Orr K, Howe HS, Omran J, *et al.*** Validity of smartphone pedometer applications. *BMC Res Notes.* 2015;8(1). doi:10.1186/s13104-015-1705-8.

10. **Bergman RJ, Spellman JW, Hall ME, Bergman SM.** Is there a valid app for that? Validity of a free pedometer iPhone application. *J Phys Act Health.* 2012;9(5):670-676.

11. **Nolan M, Mitchell JR, Doyle-Baker PK.** Validity of the Apple iPhone® /iPod Touch® as an accelerometer-based physical activity monitor: a proof-of-concept study. *J Phys Act Health.* 2014;11(4):759-769. doi:10.1123/jpah.2011-0336.