Validation of gait characteristics extracted from raw accelerometry during walking against measures of physical function, mobility, fatigability, and fitness

The Journals of Gerontology: Medical Sciences

Jacek K. Urbanek¹, Vadim Zipunnikov¹, Tamara Harris², Ciprian Crainiceanu¹, Jaroslaw Harezlak³, Nancy W. Glynn⁴

1) Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, 615 N. Wolfe Street, Baltimore, MD
2) Laboratory of Epidemiology, Demography, and Biometry, National Institute on Aging
3) Department of Epidemiology and Biostatistics, School of Public Health, Indiana University, Bloomington, IN
4) Center for Aging and Population Health, Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh PA

Address correspondence to Jacek K. Urbanek:
Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, 615 N. Wolfe Street, Baltimore, MD, email: jurbane2@jhu.edu
ABSTRACT

Background. Wearable accelerometry devices allow collection of high-density activity data in large epidemiological studies both “in-the-lab” as well as “in-the-wild” (free-living). Such data can be used to detect and identify periods of sustained harmonic walking. This report aims to establish whether the micro- and macro-features of walking identified in the laboratory and free-living environments are associated with measures of physical function, mobility, fatigability, and fitness.

Methods. Fifty-one older adults (median age 77.5) enrolled in the Developmental Epidemiologic Cohort Study in Pittsburgh, Pennsylvania were included in the analyses. The study included an “in-the-lab” component as well as 7 days of monitoring “in-the-wild”. Participants were equipped with hip-worn Actigraph GT3X+ activity monitors, which collect high-density raw accelerometry data. We applied a walking identification algorithm to the data and defined features of walking, such as participant-specific walking acceleration and cadence. The association between these walking features and physical function, mobility, fatigability, and fitness was quantified using linear regression analysis.

Results. Micro-scale features of walking (acceleration and cadence) estimated from “in-the-lab” and “in-the-wild” data were associated with measures of physical function, mobility, fatigability, and fitness. “In-the-lab” median walking acceleration was strongly inversely associated with physical function, mobility, fatigability and fitness. Additionally, “in-the-wild” daily walking time was inversely associated with usual- and fast-paced 400m walking time.

Conclusions. The proposed accelerometry-derived walking features are significantly associated with measures of physical function, mobility, fatigability, and fitness, which provides evidence of convergent validity.
**Keywords:** Wearable accelerometers, actigraphy, physical performance, gait, walking, observational studies
INTRODUCTION

The use of accelerometry has become popular in aging studies (1, 2, 3). Features derived from data collected by wearable accelerometers “in-the-lab” (or clinic) environment are often used as practical measures of mobility (4) or fatigability (5). Advanced data processing algorithms allow the analysis of specific characteristics of the human gait, such as cadence and asymmetry (6). While “in-the-lab” experiments have significant potential for medical and epidemiological studies, there is an increased interest in the data collected “in-the-wild”, that is, in the free-living environment. Several large-scale studies, including the National Health and Nutrition Examination Survey (NHANES) (7), the Baltimore Longitudinal Study on Aging (8) and the Women’s Health Initiative (WHI) (9), have either collected or are in the process of collecting “in-the-wild” activity data. As levels of activity observed “in-the-wild” have been shown to be highly correlated with both health and aging outcomes (10, 11), there is an increased interest in evaluating additional measures of activity, especially using the raw, high resolution data.

Early research on physical activity focused on aggregated measures of activity in one-minute epochs. This was due to both hardware and software limitations. Thus, most published evidence is based on a particular type of aggregation of information at the minute level. This raises questions about whether accumulating data differently can extract additional layers of information. Current accelerometry data are routinely recorded at the sub-second level, usually between 10 and 100 observations per second. For example, the most recent NHANES cohort (2013 – 2014) and the WHI study collect accelerometry data at the sub-second level. Such data offer the promise for more detailed information that can be extracted using specialized analytic approaches. Some promising early studies indicate that types of activity
can be predicted at the sub-minute level (12). In this paper, we focus on estimating gait characteristics based on “in-the-lab” and “in-the-wild” sub-second accelerometry data. Gait has been shown to be associated with longevity (13), obesity (14) and progression of Parkinson disease (15).

We use a previously developed algorithm for detection of sustained harmonic walking (SHW) (16) to extract gait characteristics in every period identified as SHW. Our aim is to examine the associations between accelerometry-derived gait characteristics and common measures of physical function, mobility, fatigability, and fitness to establish convergent validity of our novel gait characteristics. We hypothesize that: 1) features of SHW obtained from high-density accelerometry “in-the-lab” data are associated with measures of physical function, mobility, fatigability, and fitness; and 2) features of the SHW obtained from high-density accelerometry “in-the-wild” data are also associated with the same physical performance measures.

METHODS

Study participants

Eighty-nine community-dwelling older adults were recruited from the Pittsburgh, Pennsylvania area for the National Institute on Aging, Aging Research Evaluating Accelerometry (AREA) project, part of the Developmental Epidemiologic Cohort Study (DECOS) (17). AREA was a methodological initiative designed to examine the impact of accelerometry wear location on assessment of physical activity and sedentary behavior among 89 older adults enrolled between March and May of 2010. This report includes data from 51 participants (25 men and 26 women) who were available to us and had complete “in-the-lab”
(N=46) or “in-the-wild” (N=48) accelerometry data. Individuals were excluded from the study if they suffered from any of the following conditions: hip fracture, stroke in the past 12 months, cerebral hemorrhage in the past 6 months, heart attack, angioplasty, heart surgery in the past 3 months, chest pain during walking in the past 30 days, current treatment for shortness of breath or a lung condition, usual aching, stiffness, or pain in their lower limbs and joints and bilateral difficulty bending or straightening the knees fully (18).

All participants were equipped with Actigraph GT3X+ accelerometers placed on the right hip during both the “in-the-lab” and “in-the-wild” experiments. Devices collected raw accelerometry data along three orthogonal axes with the sampling frequency of 80 observations-per-second (80Hz). During the “in-the-lab” experiments participants completed both a fast-paced and a usual-paced 400m walk on two separate clinic visits. The course was set-up in a long, secured hallway with markers at both ends spaced by 20 meters. For safety purposes participants wore a heart rate monitor (Polar Chest Transmitter, Warminster, PA). The 400m walk test had two parts: a warm-up at usual pace (2 laps) and a 400m fast walk test (10 laps). Participants were excluded if, after the warm-up, their resting heart rate was over 110 or under 40 beats per minute, or if they had systolic blood pressure higher than 200 mmHg or diastolic blood pressure higher than 110 mmHg. For the fast-paced 400m walk, participants were told to walk as quickly as possible without running at a pace they could maintain for ten laps. The usual-paced 400-m walk was administered during the second clinic visit. This test was administered identically to the fast-paced 400m walk, with the exception that participants were instructed to walk at their usual, normal pace during the 400m walk portion (17). All participants were able to complete the entire sequence of activities, which
included several non-walking activities as well as resting periods. Trained research personnel administered the tests and labeled periods that corresponded to walking.

For “in-the-wild” experiment, participants were equipped with the accelerometer for seven consecutive days and were told to maintain their normal, unsupervised, free-living activities. They were instructed to take off the activity monitor only during sleep. Periods of SHW were labeled using a walking algorithm detection based on the tri-axial accelerometry data (16).

**Walking-related features**

The SHW detection algorithm (16) returns the temporal location and duration of walking bouts based on the raw accelerometry data. Additionally, the vector magnitude count (VMC) and cadence (expressed in steps-per-minute) are estimated for each walking bout. VMC is defined as the mean absolute deviation of the acceleration signal produced during SHW and is expressed in standard gravity (g=9.81 m/s²) units (16). For each estimated walking bout a set of walking features was computed including duration, VMC and cadence, for both “in-the-lab” and “in-the-wild” data. To ensure the stability of these measures, only SHW bouts longer than or equal to 20 seconds are used (30% of all identified bouts of walking).

**“In-the-lab” walking features.** For data collected during the controlled laboratory experiment we derived the median acceleration and the median cadence during SHW. The median acceleration is defined as the median of all VMCs during SHW of one individual. The median cadence is defined as the median of estimated cadence during SHW bouts. Cadence is defined as the number of steps per minute.
“In-the-wild” walking features. For data collected “in-the-wild” we derive the same micro-scale gait dynamics: median acceleration and median cadence. Additionally, we quantified the average daily walking time as the total estimated time of SHW (expressed in minutes) divided by the number of days. The average daily acceleration was obtained by dividing the total VMC during the monitoring period by the number of monitoring days. These variables depend on the total estimated SHW time and reflect the overall macro-scale activity of the individuals.

Outcomes of interest

We have focused on several outcomes of interest representing different aspects of physical performance of older adults. Physical function was measured as the time in seconds to complete five chair stands (19). Mobility was assessed using usual gait speed from a 6m walk (13) and time to finish a usual-paced 400m walk. Perceived fatigability was measured using the summary scores on the 10-item Pittsburgh Fatigability Scale for physical fatigability (PFS\textsubscript{10p}) (19). Lastly, we used time to complete a fast-paced 400m walk as a surrogate measure of aerobic fitness (18).

Statistical methods

The primary analysis consisted of determining predictive factors for each of the measures of physical function (chair stands), mobility (gait speed and usual-paced 400 meter walk time), physical fatigability, and fitness (fast-paced 400 meter walk time). We considered models where predictors were activity-derived factors, walking-derived factors and demographic information including age, sex, height and BMI. Separate models were used for: 1) “in-the-lab”
walking-derived factors: median walking acceleration and median cadence; and 2) “in-the-wild” activity-derived factors: daily acceleration, daily walking time, median walking acceleration, and median cadence. Statistical modeling was performed using linear regression and prediction performance was assessed by the adjusted R-squared for the best models chosen using the Akaike’s Information Criterion (AIC) (21).

RESULTS

Demographic characteristics of the study participants are summarized in Table 1. The median age was 77.5 years and the median BMI was 25.9 kg/m². All participants were in good overall physical health and reported no current history of medical conditions that could affect gait. All summary statistics are reported using medians and interquartile ranges. The median Pittsburgh Fatigability Scale physical fatigability score was 16.0 (12.0, 20.8) and the median time needed to complete the five-chair-stand test was 12.11 seconds (10.17, 13.90). Median usual gait speed over 6m was 1.12 (1.03, 1.24). The median time to complete the usual-pace 400 meter walk was 379.4 seconds (352.7, 414.8) while for fast-pace 400 meter walk it was 329.4 seconds (278.9, 355.5). Medians and interquartile ranges for the micro- and macro-scale walking features from the “in-the-lab” and “in-the-wild” can be found in Table 2. Estimated normalized regression coefficients and their corresponding p-values (in brackets) are shown in Table 3.

“In-the-lab” walking features (N = 46)

The median acceleration produced during fast-paced 400-meter walk had a strong negative statistical association with physical function (p < 0.001) and fatigability (p = 0.003) as well as a strong positive association with mobility (p = 0.003 for chair stands and p < 0.001 for usual-
paced 400m walk) and fitness (p < 0.001). A statistically significant positive association was found between the median cadence and mobility (p = 0.022 for chair stands and p = 0.034 for usual-paced 400m walk). The participant-specific median cadence was also negatively statistically associated with fitness (p = 0.049 for fast-paced 400m walk time).

“In-the-wild” walking features (N = 48)
Median walking acceleration was strongly negatively statistically associated with physical function (p < 0.001) and fatigability (p < 0.001) (Table 3). Median cadence was strongly statistically associated with both measures of mobility (p = 0.002 for gait speed and p = 0.034 for usual-paced 400m walk time). Additionally, there was a statistically significant negative association between daily walking time and fast-paced 400m walk time (p = 0.004). This association did not extend to daily acceleration. The median walking acceleration was also negatively statistically associated with fast-paced 400m walk time, an indicator of aerobic fitness in older adults (p = 0.002).

DISCUSSION
Using recently developed methodology for identification of bouts of walking (16), we derived measures representing both overall volume of daily walking in a free-living environment (daily walking time) and gait properties (median walking acceleration, and median cadence) for both “in-the-lab” and “in-the-wild” environment. We have found that among older adults accelerometry-derived features of SHW are significantly associated with measures of physical function, mobility, fatigability, and fitness, demonstrating the potential of the raw accelerometry data to be used as a novel source of information characterizing physical function, mobility, fatigability, and fitness.
The findings presented in this paper complement previously published results for controlled “in-the-lab” experiments. For example, a significant positive association between amplitude of walking acceleration and gait speed has been previously reported for “in-the-lab” experiments (21, 22). However, this work is novel as it demonstrated an association between median cadence and gait speed “in-the-wild”.

Micro-scale features of walking (acceleration and cadence) were associated with physical function, mobility, fatigability, and fitness. In addition, we found that “in-the-lab” median walking acceleration was associated with gait speed. Thus, median walking acceleration complements median cadence as a predictor of gait speed. These results indicate that both “in-the-lab” and “in-the-wild” experiments contain independent information about the physical performance of older adults.

Data collected “in-the-lab” is generally the main source of information in modern epidemiological studies for physical performance. Our results indicate that micro- and macro-scale gait parameters can be extracted and quantified from data collected in modern accelerometry studies both “in-the-lab” and “in-the-wild”. The amount of data collected “in-the-wild” is typically much larger, as it can be continuously measured for weeks or months at a time. Here, we propose to interpret features of SHW observed “in-the-wild” as an objective and unbiased measures of physical function, mobility, fatigability, and fitness. Indeed, even the best-designed “in-the-lab” experiment cannot fully capture the natural free-living environment conditions. By collecting data “in-the-wild” we avoid potential biases introduced by one’s tendency to under- or over-perform during supervised in-lab experiments. Moreover, data collected for extended periods in multiple bouts of SHW provides a broader image of micro- and macro-scale gait parameters.
An important pitfall with free-living data is that it can be affected by many uncontrollable factors, including type of walking surface, elevation, and environmental conditions. Deciding whether such factors are confounding the signal or are important components of it remains unresolved.

The methods used for characterizing “in-the-wild” SHW are designed for highly heterogeneous data collected without supervision. Therefore, objective measures of physical function, mobility, fatigability, and fitness can be computed and reported directly (on-line) by wearable devices or after data collection (off-line).

Our future research will focus on methodology that could serve as an alternative for “in-the-lab” measures of physical function, mobility, fatigability, and fitness using features of physical activity estimated using only “in-the-wild” data. Analysis of sub-second level data collected by wearable monitors “in-the-wild” is likely to complement or replace traditional “in-the-lab” tests in future studies.

Acknowledgements

This research was supported by Pittsburgh Claude D. Pepper Older Americans Independence Center, Research Registry, and Developmental Pilot Grant (PI: Glynn) – NIH P30 AG024826 and NIH P30 AG024827. National Institute on Aging Professional Services Contract HHSN271201100605P. NIA Aging Training Grant (PI: AB Newman) T32-AG-00181. The project was supported, in part, by the Intramural Research Program of the National Institute on Aging. This research was supported by the NIH grant RO1 NS085211 from the National Institute of Neurological Disorders and Stroke, by the NIH grant RO1 MH095836 from the National Institute of Mental Health and by NIH grant R01 HL123407 from National Hearth,
Lung and Blood Institute. Dr Harezlak’s research was supported in part by the NIMH grant R01 MH108467.
1. Copeland JL, Esliger DW. Accelerometer assessment of physical activity in active, healthy older adults. https://dspace.lboro.ac.uk/2134/15450. 2009.
2. Gardner AW, Poehlman ET. Assessment of free-living daily physical activity in older claudicants: validation against the doubly labeled water technique. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 1998;4(53):M275-M280.
3. Pruitt LA, Glynn NW, King AC, et al. Use of accelerometry to measure physical activity in older adults at risk for mobility disability. Journal of aging and physical activity;16(4):416.
4. Webber SC, Porter MM. Monitoring mobility in older adults using global positioning system (GPS) watches and accelerometers: a feasibility study. J Aging Phys Act. 2009;17(4):455-467.
5. Murphy SL, Smith DM. Ecological measurement of fatigue and fatigability in older adults with osteoarthritis. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 2010;65(2):184-189.
6. Moe-Nilssen R, Helbostad JL. Estimation of gait cycle characteristics by trunk accelerometry. Journal of biomechanics. 2004;37(1).
7. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Medicine and science in sports and exercise. Medicine and science in sports and exercise. 2008;40(1):181.
8. Talbot LA, Morrell CH, Fleg JL, Metter EJ. Changes in leisure time physical activity and risk of all-cause mortality in men and women: the Baltimore Longitudinal Study of Aging. Preventive medicine. 2007;45(2):169-176.
9. July 2015. Available at: https://www.whi.org.
10. Schrack JA, Zipunnikov V, Goldsmith J, et al. Assessing the “physical cliff”: detailed quantification of age-related differences in daily patterns of physical activity. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 2013;glt199.
11. Xiao L, Huang L, Schrack JA, Ferrucci L, Zipunnikov V, Crainiceanu CM. Quantifying the lifetime circadian rhythm of physical activity: a covariate-dependent functional approach. Biostatistics. 2015;16(2):352-367.
12. Xiao L, He B, Koster A, et al. Movement Prediction Using Accelerometers in a Human Population. arXiv preprint arXiv:1404.4601. 2014.
13. Studenski S, Perera S, Patel K, et al. Gait Speed and Survival in Older Adults. Jama. 211;305(1):50-58.
14. Browning RC, Baker EA, Herron JA, Kram R. Effects of obesity and sex on the energetic cost and preferred speed of walking. Journal of Applied Physiology. 2006;100(2).
15. Hausdorff JM, Cudkowicz ME, Firtion R, Wei JY, Goldberger AL. Gait variability and basal ganglia disorders: Stride-to-stride variations of gait cycle timing in parkinson’s disease and Huntington’s disease. Movement disorders. Movement disorders. 1998;13(3):428-437.
16. Urbanek JK, Zipunnikov V, Harris T, et al. Prediction of sustained harmonic walking in the free-living environment using raw accelerometry data. arXiv preprint arXiv:1505.04066. 2015.
17. Lange-Maia BS, Newman AB, Strotmeyer ES, Harris TB, Caserotti P, Glynn NW. Performance on fast-and usual-paced 400-m walk tests in older adults: are they comparable? Aging clinical and experimental research. 2015;27(3):309-314.

18. Lange-Maia BS, Strotmeyer ES, Harris TB, et al. Physical Activity and Change in Long Distance Corridor Walk Performance in the Health, Aging, and Body Composition Study. Journal of the American Geriatrics Society. July 2015;63(7):1348-1354.

19. Glynn NW, Santanasto AJ, Simonsick EM, et al. The Pittsburgh Fatigability Scale for older adults: development and validation. Journal of the American Geriatrics Society. 2015;63(1):130-135.

20. Akaike H. A new look at the statistical model identification. Automatic Control, IEEE Transactions on. 1974;19(6):716-723.

21. Zijlstra W, Hof AL. Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. Gait & posture. 2003;18(2):1-10.

22. Sabri K, El Badaoui M, Guillet F, Belli A, Millet G, Morin JB. Cyclostationary modeling of ground reaction force signals. Signal Processing. 2010;90(4):1146-1152.
Table 1. Characteristics of study population (N = 52)

| Variable               | Median (Q1, Q3) or N (%) |
|------------------------|----------------------------|
| Age [yr.]              | 77.5 (74.0, 82.0)          |
| Sex (Male)             | 25 (49%)                   |
| BMI                    | 25.9 (23.6, 30.0)          |
| Height [cm]            | 166.3 (159.5, 171.9)       |
| Race:                  |                            |
| White                  | 46 (90%)                   |
| Black                  | 4 (8%)                     |
| Asian                  | 1 (2%)                     |
| Education:             |                            |
| High School            | 9 (18%)                    |
| College                | 25 (49%)                   |
| Graduate               | 17 (33%)                   |
| SPPB score             | 11.0 (10.0, 12.0)          |
| Physical fatigability (psf10p) | 14.0 (11.5, 20.5)        |
| Chair stands           | 12.1 (10.2, 13.9)          |
| Usual gait speed       | 1.11 (1.03, 1.23)          |
| Usual 400m walk time (sec) | 379.4 (352.7, 414.8)    |
| Fast 400m walk time (sec) | 329.4 (278.9, 355.5)   |

Table 2. Medians and values of Q1 and Q3 (in the brackets) for the micro- and macro-scale walking features.

| Variable                                | Median (Q1, Q3) or N (%) |
|-----------------------------------------|----------------------------|
| "in-the-wild"                           |                            |
| Daily Acceleration [g]                  | 71273 (Q1 = 51387, Q3 = 91565) |
| Median Cadence [steps/min.]             | 101.50 (Q1 = 93.56, Q3 = 106.25) |
| Daily Walking Time [min.]               | 60.334 (Q1 = 38.712, Q3 = 81.244) |
| Median Walking Acceleration [g]         | 0.12982 (Q1 = 0.09590, Q3 = 0.14784) |
| "in-the-lab"                            |                            |
| Median Walking Acceleration [g]         | 0.2254 (Q1 = 0.1831, Q3 = 0.3014) |
| Median Cadence [steps/min.]             | 124.0 (Q1 = 117.0, Q3 =129.5)  |
Table. 3. Estimated normalized regression coefficients and the corresponding p-values (in brackets) of the best-fitted models. Values on the gray-shaded lines denote the estimates of the adjusted $R^2$.

| Physical Function (5 Chair stands) | Mobility (4m Gait speed) | Mobility (Usual-paced 400m walk) | Physical Fatigability (PFS10p) | Fitness (Fast-paced 400m walk) |
|-----------------------------------|--------------------------|----------------------------------|-------------------------------|-------------------------------|
| “In-the-lab”                      |                          |                                  |                               |                               |
| Median walking acceleration        | -0.531 (-<0.001)         | 0.381 (0.003)                    | -0.636 (-<0.001)              | -0.422 (-0.003)               |
| Median cadence                    | -0.268 (0.055)           | 0.298 (0.022)                    | -0.234 (0.034)                | -0.215 (0.049)                |
| Age                               |                          |                                  |                               |                               |
| Height                            |                          | 0.479 (-<0.001)                  | -0.511 (-<0.001)              |                               |
| BMI                               |                          |                                  |                               |                               |
| Sex                               |                          | -0.365 (0.011)                   |                               |                               |
| “In-the-wild”                     | 0.23                     | 0.27                             | 0.38                          | 0.30                          |
| Daily acceleration                |                          |                                  |                               |                               |
| Daily walking time                |                          | -0.255 (0.048)                   | -0.318 (0.004)                |                               |
| Median walking acceleration        | -0.433 (-<0.001)         | -0.799 (-<0.001)                 | -0.551 (-<0.001)              | -0.378 (0.002)                |
| Median cadence                    | 0.378 (0.002)            | -0.362 (0.034)                   |                               |                               |
| Age                               |                          |                                  |                               | 0.344 (0.013)                 |
| Height                            | 0.3324 (0.005)           | -0.548 (0.004)                   |                               |                               |
| BMI                               |                          | 0.263 (0.015)                    |                               |                               |
| Sex                               |                          | -0.63163 (0.003)                 |                               |                               |