Journal for the Measurement of Physical Behaviour, 2022, 5, 145-155
https://doi.org/10.1123/jmpb.2021-0031

Human Kinetics
ORIGINAL RESEARCH

Calibration of an Accelerometer Activity Index Among Older Women and Its Association With Cardiometabolic Risk Factors

Guangxing Wang,[1] Sixuan Wu,[2] Kelly R. Evenson,[3] Ilsuk Kang,[1] Michael J. LaMonte,[4] John Bellettiere,[5] I-Min Lee,[6,7] Annie Green Howard,[8,9] Andrea Z. LaCroix,[5] and Chongzhi Di[1]

[1]Division of Public Health Sciences, Fred Hutchinson Cancer Center, Seattle, WA, USA; [2]Inspur USA Inc., Bellevue, WA, USA; [3]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; [4]Department of Epidemiology and Environmental Health, School of Public Health and Health Professions, University at Buffalo—SUNY, Buffalo, NY, USA; [5]Division of Epidemiology, Herbert Wertheim School of Public Health and Human Longevity Science, University of California, San Diego, CA, USA; [6]Division of Preventive Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; [7]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA; [8]Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; [9]Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Purpose: Traditional summary metrics provided by accelerometer device manufacturers, known as counts, are proprietary and manufacturer specific, making it difficult to compare studies using different devices. Alternative summary metrics based on raw accelerometry data have been introduced in recent years. However, they were often not calibrated on ground truth measures of activity-related energy expenditure for direct translation into continuous activity intensity levels. Our purpose is to calibrate, derive, and validate thresholds among women 60 years and older based on a recently proposed transparent raw data-based accelerometer activity index (AAI) and to demonstrate its application in association with cardiometabolic risk factors. ***Methods:*** We first built calibration equations for estimating metabolic equivalents continuously using AAI and personal characteristics using internal calibration data ($N = 199$). We then derived AAI cutpoints to classify epochs into sedentary behavior and physical activity intensity categories. The AAI cutpoints were applied to 4,655 data units in the main study. We then utilized linear models to investigate associations of AAI sedentary behavior and physical activity intensity with cardiometabolic risk factors. ***Results:*** We found that AAI demonstrated great predictive accuracy for estimating metabolic equivalents ($R^2 = .74$). AAI-Based physical activity measures were associated in the expected directions with body mass index, blood glucose, and high-density lipoprotein cholesterol. ***Conclusion:*** The calibration framework for AAI and the cutpoints derived for women older than 60 years can be applied to ongoing epidemiologic studies to more accurately define sedentary behavior and physical activity intensity exposures, which could improve accuracy of estimated associations with health outcomes.

Keywords: accelerometry, older adults, physical activity, sedentary behavior, validation

Accelerometers have been widely used to measure physical activity (PA) in biomedical studies in past decades (Bai et al., 2016; Yang & Hsu, 2010). Traditionally, the common data output consists of summary measures over user-defined epochs (e.g., 1 min; Chen et al., 2012). For example, both ActiGraph (ActiGraph) and Actical (Philips Respironics) software use proprietary algorithms to calculate counts per epoch (John & Freedson, 2012), which is an aggregated measure of the acceleration magnitude over a given time epoch. Although easy to use, the proprietary nature of counts makes comparison between studies using different devices difficult because the counts defined by different manufacturers can have very different meanings (Rowlands, 2018).

With technological advances of recent years, several research grade accelerometers allow direct measurement and access to high-resolution raw acceleration data (10–100 Hz), which contain much richer information than proprietary counts on epoch-level resolutions. This enables researchers to use well-defined, open-source, reproducible summaries of the data to compare and combine studies that collect raw accelerometry data (Karas et al., 2019). To overcome limitations of counts, efforts have been made to establish summary metrics from the raw data with transparent formulas. A few raw accelerometry data-based metrics have been proposed, such as an accelerometer activity index (AAI; Bai et al., 2016), Euclidean norm minus one (van Hees et al., 2013), mean amplitude deviation (Aittasalo et al., 2015), and a monitor-independent movement summary (John et al., 2019). In general, all these newly proposed metrics aim to quantify and summarize the magnitude of acceleration during a given epoch. For these metrics to be useful and interpretable, it is important to calibrate them against ground truth PA energy expenditure (e.g., intensity of PA). The gold standard measure of PA energy expenditure, or intensity, is oxygen uptake (VO_2) directly measured using calorimetry procedures during performance of sedentary behaviors (SB) and PA tasks (LaMonte et al., 2006). Calibration studies are useful in translating PA metrics into intensity-specific categories relevant to specified populations (Jago et al., 2007). For some health outcomes, PA intensity may be a more relevant exposure than the amount of PA performed (Kesaniemi et al., 2001). To the best of our knowledge, little work on calibration of raw accelerometry data-based metrics exists, especially among older adults. It is particularly important to calibrate the intensity-specific cutpoints among adults 60 years and older because the cutpoints for this group may differ dramatically from those derived in samples of younger adults (Evenson et al., 2015) given the increase in absolute energy cost of movement associated with aging (Ortega & Farley, 2007).

Di (cdi@fredhutch.org) is corresponding author, https://orcid.org/0000-0002-1371-581X

We conducted a novel analysis based on an AAI, a recently proposed transparent metric to summarize high-dimensional raw accelerometry data, in women aged 60 years and older. Calibration equations were derived for metabolic equivalents (METs) continuously using AAI; the AAI cutpoints were obtained to classify epochs into SB and distinct PA intensity categories using data from a calibration study on older women. We then utilized the calibrated AAI to investigate associations of SB and PA measures with cardiometabolic risk factors. To investigate simultaneous associations of multiple intensity categories, we also adopted isotemporal models to quantify the potential substitutional effects of reallocating time between two activity intensity categories on cardiometabolic risk factors.

Materials and Methods

Study Design and Participants

The Women's Health Initiative (WHI), sponsored by the National Heart, Long and Blood Institute, is an ongoing study of the determinants of morbidity and mortality in aging postmenopausal women. WHI enrolled 161,808 postmenopausal women, aged 50–79 years, into one of three randomized clinical trials or a prospective observational study across 40 U.S. study sites from 1993 to 1998 (Anderson et al., 2003). During the WHI 2010–2015 extension study, home examinations were completed in a subcohort of 7,875 women also enrolled into the ancillary Long Life Study (2012–2013). The Objective Physical Activity and Cardiovascular Health (OPACH) Study (LaCroix et al., 2017) was ancillary to the Long Life Study and consisted of ambulatory community-dwelling women aged 63–99 years. Women were recruited to wear an ActiGraph GT3X+ accelerometer to measure free-living PA and SB from May 2012 to April 2014 and relate these measures to cardiovascular disease incidence during follow-up. Overall, 7,048 women received the accelerometer and a sleep log, among whom 6,721 women (95.4%) returned accelerometers and 6,489 (92.0%) had data for at least one day. Among them, 6,078 women had valid raw data for computing AAI, and 5,870 had data for at least four adherent days using the common definition of at least 10 hr of accelerometer wear time while awake. Study protocols were approved by the Fred Hutchinson Cancer Research Center (WHI Coordinating Center). All women provided informed consent in writing or by phone.

Free-Living Accelerometry Data Collection

OPACH women wore an ActiGraph GT3X+ triaxial accelerometer over the right hip 24 hr per day for up to seven consecutive days except when they risked submerging the accelerometer in water (e.g., during bathing or swimming). Sleep time was identified using self-reported in-bed and out-of-bed times from sleep diaries that were filled out each night of accelerometer wear (Rillamas-Sun et al., 2015). If their sleep log data were missing, their in-bed and out-of-bed times were imputed using person-specific means, if available, or the mean over the full OPACH sample. Actigraph's proprietry ActiLife (version 6) software was used to process raw data (30 Hz) into counts per 15-s epochs. Vector magnitude counts were derived by taking the square root of the sum of the three axes squared to capture movement in all three axes. Nonwear periods were identified using the Choi algorithm on vector magnitude counts (Choi et al., 2011, 2012). We then applied count-based intensity-specific cutpoints that were determined using the same sample as the present study (Evenson et al., 2015), that is, moderate to vigorous PA (MVPA; >518 counts per 15 s), high light PA (HLPA; >225 counts per 15 s and ≤518 counts per 15 s), low light PA (LLPA; >18 counts per 15 s and ≤225 counts per 15 s), and SB (≤18 counts per 15 s).

Health Outcomes

Questionnaires ascertained participant age, race/ethnicity (Black, White, or Hispanic/Latina), and educational attainment (high school equivalent or lower, some college, or college graduate). During the Long Life Study home examinations, trained study staff measured height (in meters) and weight (in kilograms) with a tape measure and calibrated scale, respectively, and calculated body mass index (BMI, in kilograms per square meter). Fasting (12 hr) blood samples were obtained, and cardiometabolic risk factors including serum glucose and high-density lipoprotein (HDL) were measured using standardized Clinical Laboratory Improvement Act-approved methods at the University of Minnesota (LaMonte et al., 2017).

Calibration Study

The OPACH study included one laboratory session where 200 women aged 60–91 years were invited to participate in a study to calibrate accelerometry to energy expenditure during SBs and various PA tasks differing in known energy costs. The participants were asked to perform several standardized tasks while simultaneously using a hip-worn accelerometer, a wrist-worn accelerometer, a heart rate monitor, and a portable indirect calorimeter to measure VO_2. The VO_2 measures were expressed in METs, which represent multiples of metabolic energy cost defined as the ratio of activity-related energy expenditure to resting energy expenditure (LaMonte et al., 2006). A portable breath-by-breath metabolic device, Oxycon Mobile (CareFusion), was used to measure VO_2, and a chest-worn POLAR heart rate monitor (Polar Electro) was used to measure heart rate continuously during the PA tasks. The intensity of selected tasks of the calibration study varied from sedentary (<1.5 METs) to light intensity (1.5–3.0 METs) and moderate intensity (>3.0–6.0 METs) for older women. Participants were asked to rest 2 min between activities to allow their heart rate to return to within 10 beats per minute of their resting heart rate. The duration of tasks was 7 min, except for the usual pace 400-m walk, so that participants achieved steady-rate metabolism for measurement of task-specific VO_2. The participants performed the following tasks: watching a DVD while sitting quietly (DVD), wash/dry dishes while standing (DISHES), laundry (removing towels from basket and folding) while standing (LAUNDRY), 400-m walk (WALK), assemble puzzle while sitting (PUZZLE), and dust mopping while standing (MOPPING). The Oxycon Mobile measured VO_2 was converted to average energy expenditure during each activity in METs, by dividing the resting oxygen intake by 3.0 ml/kg·min. The 3.0 ml/kg·min is a departure from the conventional value 3.5 ml/kg·min, but given the age-related decline in resting metabolic rate (Tzankoff & Norris, 1977), a lower value is a more accurate estimate of resting metabolic rate for older adults as observed by other investigators (Reidlinger et al., 2015). Using too large a value for resting metabolic rate (e.g., 3.5) in the denominator of a MET introduces a downward bias (underestimation) of the participant's activity-related MET value (Kozey et al., 2010). More details about these measurements and protocol can be found elsewhere, including the collection of descriptive characteristics, weight, and height (Evenson et al., 2015).

Statistical Analysis

Summary statistics described distributions of participant characteristics for the OPACH main study and calibration study. Using the equation in Bai et al. (2016; implemented in an R package "ActivityIndex"), we first calculated AAI per second for each participant based on 30-Hz raw data from hip-worn accelerometers. In particular, we calculated the AAI in relative scale (their Formula 2) with the sigma value (systematic noise *SD*) estimated based on raw acceleration signals when the device is not moving (0.002559424). They were then aggregated into AAI per 15 s by taking the sum of AAI per second within the 15-s epoch. The AAI was shown to have desirable mathematical properties including additivity and rotational invariance, and thus was not affected by variation in device orientations among participants. In the OPACH calibration study, we used data during Minutes 3–7 for activities lasting 7 min (DVD, DISHES, LAUNDRY, PUZZLE, and MOPPING). For the 400-m walk, we used data from Minute 3 to the end of the walk. Oxygen uptake during these activities measured by Oxycon Mobile devices was used to calculate activity-specific intensity defined in units of METs, which represent the ratio of activity energy expenditure to resting energy expenditure (we used 3.0 $\text{ml}\cdot\text{min}^{-1}\cdot\text{kg}^{-1}$, the median value measured in our sample while sitting quietly watching a DVD; Evenson et al., 2015). After deriving AAIs, histograms were plotted to show their distributions for each of the six activity types.

Linear regression was used to study the relationship between AAI and METs and to build calibration equations to predict continuous METs. Three types of transformations of AAI were explored: original, square root, and logarithm scales. When modeling the relationship between METs and AAI per 15 s, the fitted line was forced to go through the point (0,1) to reflect prior knowledge that a zero AAI per 15 s generally implies that a person stays still with little movement, so the corresponding activity intensity is exactly 1 MET. We fit two models, one with AAI only and the other allowing an additional interaction with age groups (selected based on quartiles to ensure equal group sizes). The latter yielded age-specific calibration equations. Fivefold cross-validation was used to estimate root mean square error and R^2 to evaluate the predictive performance of these models (Hastie et al., 2017).

Receiver operating characteristic (ROC) curve analysis was conducted to derive AAI cutpoints for classifying a 15-s epoch into one of the intensity categories: SB, LLPA, HLPA, and MVPA (Jago et al., 2007). We identified cutpoints by balancing the number of false positives and false negatives. The area under the ROC curve (AUC) was estimated using generalized estimating equations with logistic regression (Liang & Zeger, 1986) to account for within-person correlations. The AUC represented the predictive accuracy of accelerometer metrics to classify activity intensity categories (Pepe, 2004).

In the OPACH main study, we first calculated AAI per 15 s during wear periods while awake for each woman and then applied the derived cutpoints to estimate daily average minutes spent in each intensity category. To study the relationships between AAI-derived activity measures and BMI, glucose, and HDL, we fit linear regression models. Two types of linear regression models, the single activity linear regression model and the isotemporal substitution model, were used to examine the associations between time spent in PA and SB per day and BMI, glucose, and HDL (Mekary et al., 2009; Stamatakis et al., 2015). For the glucose and HDL cholesterol outcomes, both models adjusted for average awake wear time, age, race and ethnicity, education, and BMI groups. For the BMI outcome, we only adjusted for average awake wear time, age, race and ethnicity, and education. For comparison, we repeated the above association analysis with the same statistical methods using counts-based PA variables that are derived with data from the same cohort and same laboratory.

Results

Description of Samples

The analysis sample from the main study included 4,655 OPACH women who had at least four days of complete accelerometry data as well as an assessment of their cardiometabolic risk factors (Table 1). They had mean age of 78.9 (*SD* 6.7) years, and their BMI was evenly distributed among normal or underweight (33.4%), overweight (30.1%), and obese (36.5%) categories. A total of 199 of the 200 women in the OPACH calibration study had complete raw data (30 Hz) available and were our analysis sample. They had mean age of 75.4 (*SD* 7.7) years, and their BMI was also evenly distributed among normal or underweight (36.2%), overweight (30.6%), and obese (33.2%) categories.

Table 1 OPACH Main (*N* = 4,655) and Calibration (*N* = 199) Study Participant Characteristics

	OPACH main study	OPACH calibration study
	Mean (*SD*)	Mean (*SD*)
Age, years	78.9 (6.6)	75.5 (7.7)
BMI, kg/m^2	27.9 (5.7)	28.0 (6.0)
HDL cholesterol, mg/dl	60.5 (15.0)	
Glucose, mg/dl	98.2 (27.6)	
	n (%)	*n* (%)
Age, years		
60–69	461 (9.9)	43 (21.6)
70–79	1,813 (38.9)	88 (44.2)
80+	2,381 (51.1)	68 (34.2)
Race/ethnicity		
White	2,441 (52.4)	100 (50.3)
Black	1,411 (30.3)	64 (32.2)
Hispanic	803 (17.3)	35 (17.6)
Education		
High school or less	940 (20.2)	29 (14.6)
Some college	1,778 (38.2)	68 (34.2)
College graduate	1,913 (41.1)	102 (51.3)
BMI, kg/m^2		
Underweight	73 (1.6)	3 (1.5)
Normal	1,430 (30.7)	69 (34.7)
Overweight	1,672 (35.9)	63 (31.7)
Obese	1,387 (29.8)	64 (32.2)

Note. BMI = body mass index; HDL= high-density lipoprotein; OPACH = Objective Physical Activity and Cardiovascular Health.

Calibration: Derivation of Calibration Equations for AAI

The AAIs were calculated and histograms for AAI per 15 s are shown in Figure 1. The mean (*SD*) of measured METs were 1.0 (0.2) for DVD, 1.3 (0.3) for PUZZLE, 1.8 (0.4) for DISHES, 2.0 (0.4) for LAUNDRY, 2.5 (0.6) for MOPPING, and 3.7 (0.7) for WALK, and the mean (*SD*) of corresponding AAI per 15 s were 9.2 (27.6), 78.1 (45.1), 139.7 (72.1), 203.3 (71.7), 383.7 (139.9), and 928.3 (258.0), respectively. As expected, epochs with higher intensity activities (METs) had higher AAI levels.

Linear regression models were used to build calibration equations for continuous METs using AAIs. After exploring three transformations of AAI (original, logarithm, and square root), we found that the square root transformation yielded the best fit (Figure 2) and thus used it in subsequent analysis. Table 2 shows model fitting results from two models: univariate (AAI only, Model 1) and interaction with age (Model 2). Note that the models had no intercept or main effects of age, as it is desirable to force MET = 1 when AAI = 0, regardless of age. Based on Model 1, AAI was highly predictive for METs with an R^2 of .74. This model provided a calibration equation, $\text{MET} = 1 + 0.08\sqrt{\text{AAI}}$, which implied a strong positive relationship between AAI and METs. Allowing age-specific relationships slightly improved predictive performance ($R^2 = .77$ in Model 2), as illustrated by Figure 3.

Calibration: Derivation of AAI Cutpoints

Table 3 displays ROC-derived cutpoints for classifying AAI per 15-s epochs into one of the four intensity categories: SB, LLPA, HLPA, and MVPA. The estimated AAI cutpoints were 101, 270, and 573 for SB versus LLPA, LLPA versus HLPA, and HLPA versus MVPA, respectively, with corresponding AUC values of 0.93, 0.94, and 0.97, respectively. Table 4 contains descriptive statistics of PA-related metrics derived using both AAI- and count-based cutpoints for the OPACH main study. Notably, AAI-based metrics implied an average of 9.3 min/day more SB, 8.5 min/day less LLPA, 11.5 min/day more HLPA, and 12.3 min/day less MVPA for these women compared with count-based metrics.

We also conducted ROC analysis to derive age-specific cutpoints for each of four age categories: 60–69, 70–74, 75–81, and 82+ years old (Supplementary Table S1 [available online]). The results showed that cutpoints for each intensity category decreased with age. These cutpoints were then used to calculate daily summary variables for SB, LLPA, HLPA, and MVPA. Descriptive statistics for these variables (Supplementary Table S2 [available online]) showed generally similar patterns with those based on universal cutpoints (Table 4), with around 5 min/day in mean differences of LLPA, HLPA, and MVPA and 16 min/day in mean difference of SB. Similarly, we derived and presented BMI-specific cutpoints for women in each of three categories: <25, 25–30, >30 kg/m^2 (Supplementary Table S5 [available online]).

Association With Health Outcomes

Associations of each PA measure and three cardiometabolic risk factors (BMI, glucose, and HDL) were estimated from linear regression. Table 5 shows model fitting results, where each model contained only one PA measure and was adjusted for average awake wear time (in minutes), age, BMI (except when BMI was the outcome), race and ethnicity, and education. Each coefficient represented effects of 30 min/day difference in time spent in the corresponding activity intensity category. For AAI-based measures, all associations were statistically significant. There were generally monotonic dose–response relationships between all three cardiometabolic outcomes and PA with varying intensity levels. For example, each 30 min/day incremental increase in SB was associated with 0.74 mg/dl lower HDL, while each 30 min/day incremental difference in LLPA, HLPA, and MVPA was associated with 1.00, 1.46, and 2.15 mg/dl higher HDL, respectively, demonstrating stronger associations with higher intensity PA. Comparing analysis based on two sets of measures, AAI-based analysis provided a clearer dose–response pattern across varying PA intensity levels for each health outcome. For example, count-based analysis estimated similar effects of HLPA and MVPA on HDL (1.68 and 1.61 mg/dl increase per 30 min/day increase in HLPA and MVPA, respectively), while AAI-based analysis estimated stronger association with MVPA than HLPA (2.15 and 1.46, respectively).

We also fit isotemporal substitution models to estimate the potential effect associated with reallocating 30 min/day between two activity intensity categories (Table 6). In this analysis, LLPA and HLPA were combined into a single category, light PA. For AAI-based measures, theoretically reallocating 30 min/day of SB by MVPA was significantly associated with lower BMI and glucose and a higher HDL. Substitution from SB to light PA was associated with lower BMI and glucose and a higher HDL, but the effect is not as strong as that when reallocating SB to MVPA. Reallocation between light PA and MVPA shows similar patterns. AAI-Based analysis had higher power to detect substitutional effects than count-based analysis. For example, effects of reallocating LPA to MVPA were not significantly associated with BMI and HDL in count-based analysis, but the associations were stronger and statistically significant in AAI-based analysis.

As sensitivity analysis, we repeated the regression models using AAI-derived intensity measures based on age-specific cutpoints (Supplementary Tables S3 and S4 [available online]). There were quantitative differences between these results and prior results based on universal cutpoints. However, the overall comparisons between AAI- and counts-based analyses remain similar, that is, AAI-based analysis provided a clearer dose–response pattern between PA intensity measures and outcome variables and had higher power to detect substitutional effects, especially effects of reallocating light PA to MVPA.

Discussion

This paper provides calibration of AAI, a raw accelerometry-based metric among older women. Using OPACH calibration study data, AAI from hip-worn triaxial ActiGraph GT3X+ accelerometers was calibrated to ground truth VO_2 measured during several activity tasks of daily living, yielding AAI cutpoints for various PA intensity categories. PA intensity measures derived from these cutpoints were then used to study associations between PA intensity and cardiometabolic risk factors in the OPACH main study. Our study provides a framework for calibrating raw accelerometry-based metric and deriving intensity cutpoints from such calibration, which can be used in subsequent association analysis. The activity intensity cutpoints and calibration equations derived in this study can be applied to other studies that collect raw accelerometry data in similar populations, for example, the Women's Health Study (Shiroma et al., 2013). For different populations, such as different age or gender groups, similar methods can be directly applied to derive population-specific intensity cutpoints and calibration equations if calibration study data are available. In particular, this is not

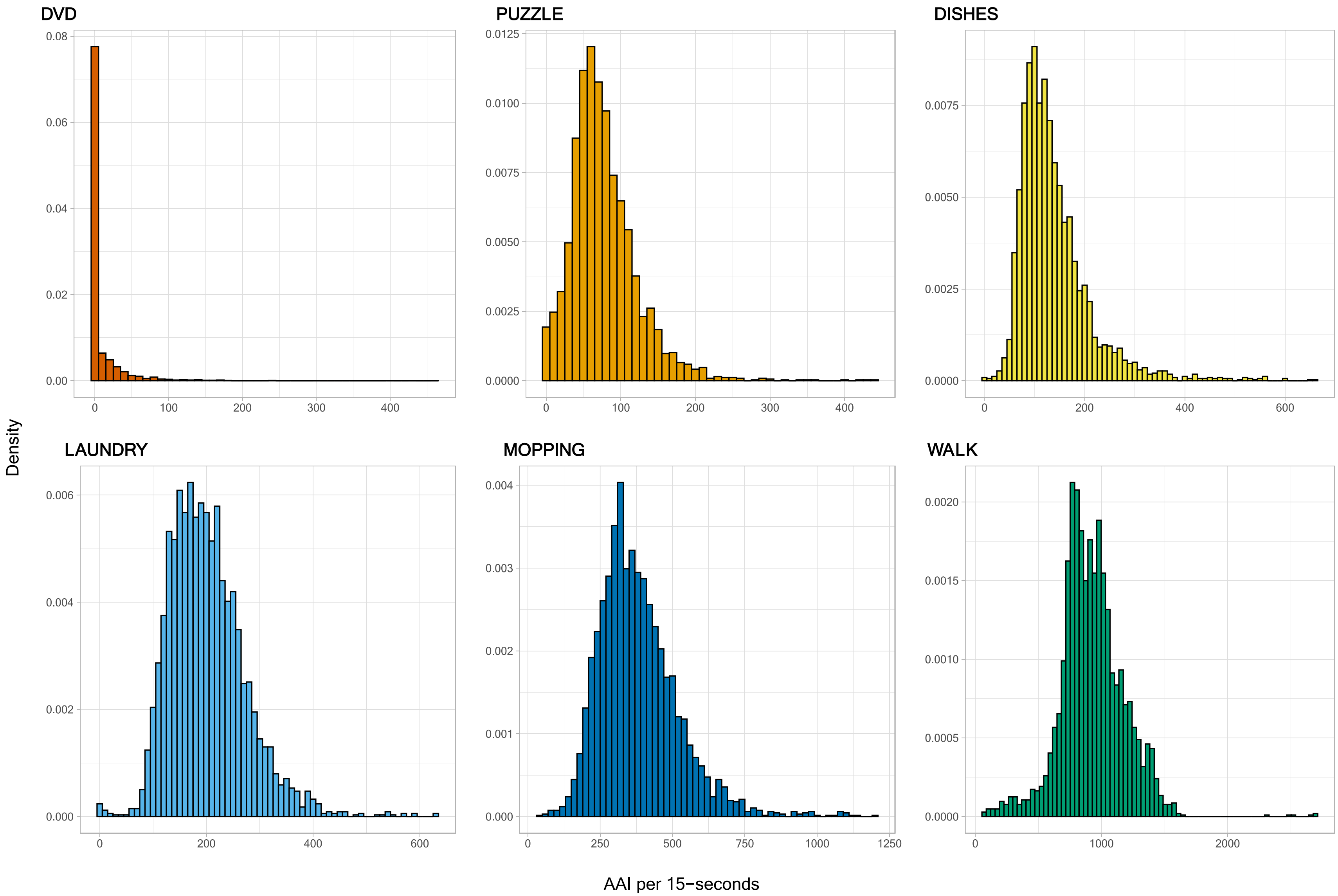

Figure 1 — Histograms of AAI per 15 s by activity type in OPACH calibration study ($N = 199$). The six types of activities are as follows: DVD, watching DVD while sitting quietly; PUZZLE, assembling puzzle while sitting; DISHES, washing dishes while standing; LAUNDRY, doing laundry while standing; MOPPING, dust mopping while standing; and WALK, 400-m walking. The *x*- and *y*-axis across the figures are different. AAI = accelerometry activity index; OPACH = Objective Physical Activity and Cardiovascular Health.

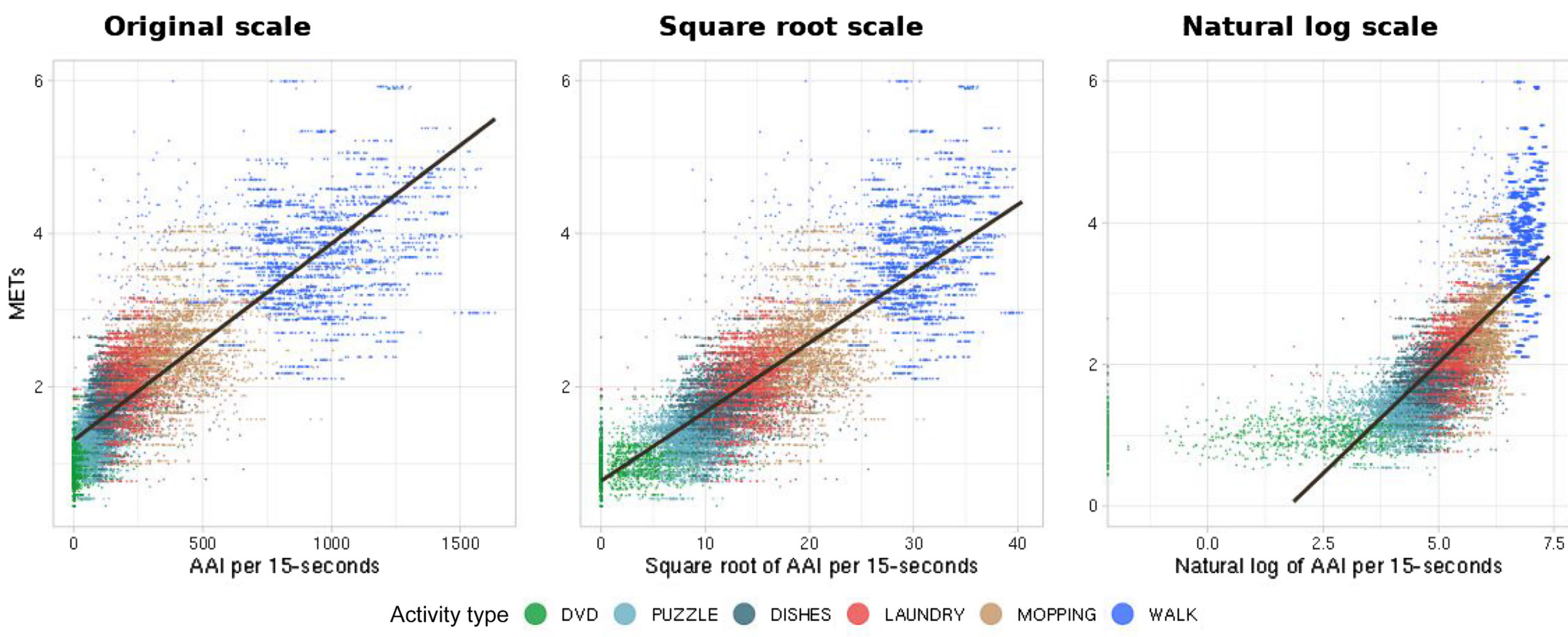

Figure 2 — Scatterplots of METs versus AAI per 15 s (original, square root, and logarithm scales) in OPACH calibration study ($N = 199$), with linear fitted lines superimposed. The six types of activities are as follows: DVD, watching DVD while sitting quietly; PUZZLE, assembling puzzle while sitting; DISHES, washing dishes while standing; LAUNDRY, doing laundry while standing; MOPPING, dust mopping while standing; and WALK, 400-m walking. METs = metabolic equivalents; AAI = accelerometry activity index; OPACH = Objective Physical Activity and Cardiovascular Health.

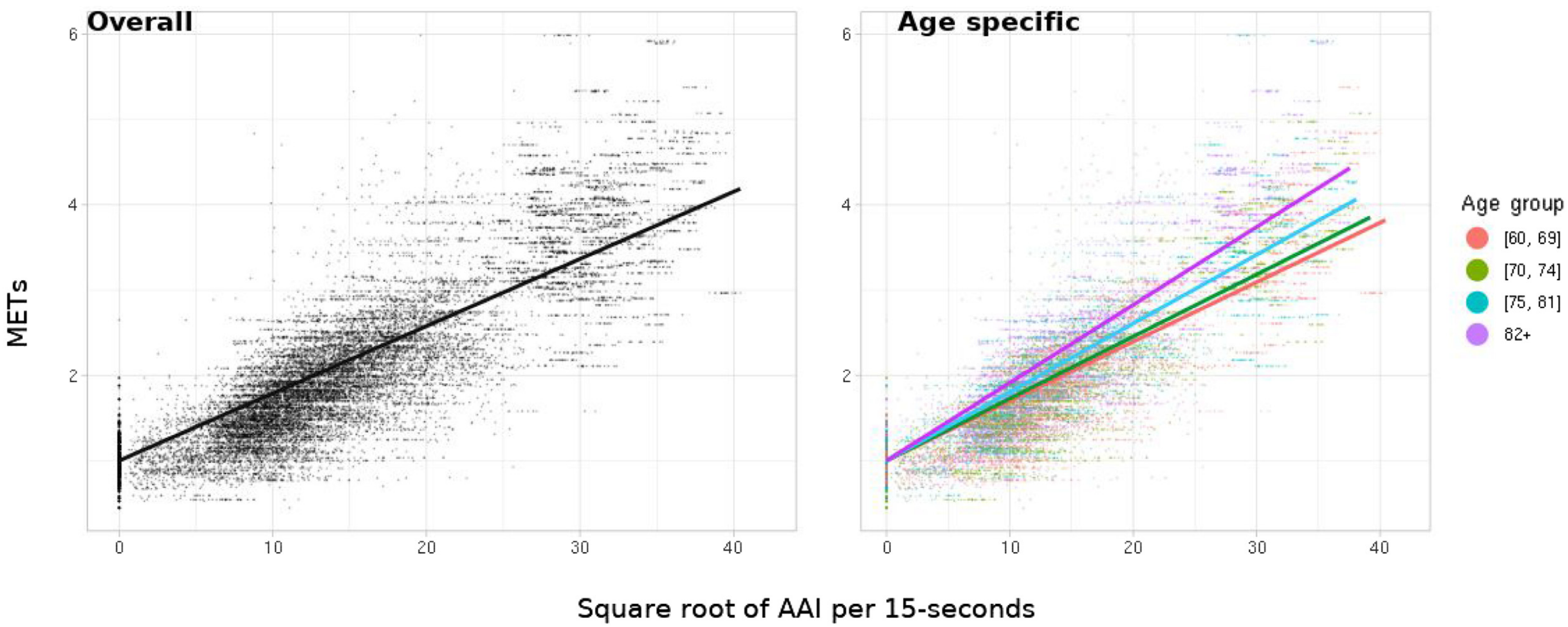

Figure 3 — Scatterplots between METs and AAI per 15 s from the OPACH calibration study ($N = 199$). The left panel illustrates fitted linear model using the full sample. The right panel added fitted lines between METs and AAI for each age category. METs = metabolic equivalents; AAI = accelerometry activity index; OPACH = Objective Physical Activity and Cardiovascular Health.

limited to AAI, and it applies to any raw accelerometry-based metrics.

We applied ROC analysis to derive AAI cutpoints for PA intensity classification. Evaluation metrics such as sensitivity, specificity, and AUC were used to demonstrate that the AAI had great predictive performance for activity intensity. For instance, AUCs for AAI-based classification were 0.93, 0.94, and 0.97 for SB versus LLPA, LLPA versus HLPA, and HLPA versus MVPA, respectively. We derived the cutpoints by balancing the number of false positives and false negatives in ROC analysis. As pointed out by Evenson et al. (2015), this approach provides roughly unbiased estimates for variables with low positive predicted values such as MVPA for older adults, for whom the prevalence of such activities tends to be relatively low.

To illustrate the usefulness of the derived AAI cutpoints, we then investigated associations of AAI-based intensity measures and three cardiometabolic risk factors: BMI, glucose, and HDL cholesterol in the OPACH main study. Based on prior reviews (2018 Physical Activity Guidelines Advisory Committee, 2018; Barone Gibbs et al., 2021; Jakicic et al., 2019; Katzmarzyk et al., 2019), we would expect MVPA to be positively associated with HDL and inversely associated with glucose and BMI, while SB would be inversely associated with HDL and positively associated with glucose and BMI. Linear regression models were used to estimate association between specific activity intensity and outcomes, while isotemporal substitution models estimated an expected effect of reallocating equivalent time units between two different intensity categories. These analyses implied that greater time spent in light PA or MVPA were associated with a more favorable cardiometabolic risk profile, as was reallocating SB to light PA or MVPA. We compared the linear regression results based on AAI to those based on count-derived measures, which were previously reported by LaMonte et al. (2017). Note that we reanalyzed counts data to ensure more meaningful comparison with AAI-based analysis as our analysis sample ($N = 4{,}655$) was slightly smaller than their original analysis ($N = 4{,}832$). The conclusions were similar for most associations, although there were a few noticeable differences. Effects of MVPA and the cardiometabolic outcomes were weaker with counts-based measures. In addition, with count-based measures, the dose–response pattern across PA intensity categories was not as clear. For example, MVPA associations were even weaker than HLPA associations with BMI and HDL. For isotemporal substitution models, analysis using count-based metrics yielded similar conclusions for reallocation of equivalent time from SB to light PA, although results were different for a few associations involving MVPA. Notably, reallocating time from light PA to MVPA was not statistically significant with BMI and HDL outcomes in counts-based measures, but was significant for AAI-based measures. Overall, AAI-based summary measures often demonstrated stronger associations, especially for MVPA, and clearer dose–response relationships across varying intensity levels compared with count-based measures.

Regarding comparison between AAI and counts, it is worth noting that AAI demonstrated better predictive accuracy for METs

Table 2 Calibration Equations for METs Based on AAI per 15-s Epochs From Hip-Worn Accelerometers; WHI OPACH Calibration Study ($N = 199$)

Variable	Calibration equation	RMSE	R^2
Model 1: Overall		.503	.74
AAI	$\text{MET} = 1 + 0.079\sqrt{\text{AAI}}$		
Model 2: Age specific		.482	.77
60–69 years old	$\text{MET} = 1 + 0.070\sqrt{\text{AAI}}$		
70–74 years old	$\text{MET} = 1 + 0.073\sqrt{\text{AAI}}$		
75–81 years old	$\text{MET} = 1 + 0.080\sqrt{\text{AAI}}$		
82+ years old	$\text{MET} = 1 + 0.091\sqrt{\text{AAI}}$		

Note. All models have p value < .001. AAI = accelerometer activity index; METs = metabolic equivalents; RMSE = root mean square errors; OPACH = Objective Physical Activity and Cardiovascular Health; WHI = Women's Health Initiative.

Table 4 Summary Statistics of Metrics Derived From AAI and Count Cutpoints for the OPACH Cohort ($N = 4{,}655$)

PA-Related metrics	Q1	Mean	*SD*	Median	Q3
Awake wear time (min/day)	846.6	893.9	78.0	900.1	947.7
Counts per 15 s	72.5	103.0	42.4	97.4	128.1
AAI per 15 s	107.3	140.7	46.6	135.4	168.5
Count-based intensity (min/day)					
SB time	487.5	553.2	99.3	555.5	619.6
LLPA time	154.0	189.3	49.8	187.2	220.0
HLPA time	74.6	99.5	35.6	97.1	121.9
MVPA time	25.5	51.9	35.0	45.0	70.5
AAI-based intensity time (min/day)					
SB time	492.4	562.5	104.1	563.9	633.1
LLPA time	142.3	180.8	53.5	178.4	216.1
HLPA time	82.8	111.0	40.5	107.1	135.1
MVPA time	18.2	39.6	28.7	33.8	54.5

Note. AAI = accelerometer activity index; OPACH = Objective Physical Activity and Cardiovascular Health; SB = sedentary behavior; LLPA = low light physical activity; HLPA = high light physical activity; MVPA = moderate and vigorous physical activity.

Table 3 Hip-Worn Accelerometer Cutpoints for AAI per 15 s Derived From ROC-Based Approach ($N = 199$); WHI OPACH Calibration Study ($N = 199$)

Activity intensity	Cutpoints (AAI per 15 s)	Sensitivity	Specificity	Sensitivity + specificity	AUC
SB vs. LLPA	101	0.79	0.88	1.67	0.92
LLPA vs. HLPA	270	0.79	0.90	1.70	0.94
HLPA vs. MVPA	587	0.82	0.97	1.79	0.97

Note. AAI = accelerometer activity index; OPACH = Objective Physical Activity and Cardiovascular Health; WHI = Women's Health Initiative; ROC = receiver operating characteristic; AUC = area under the ROC curve; SB = sedentary behavior; LLPA = low light physical activity; HLPA = high light physical activity; MVPA = moderate and vigorous physical activity.

Table 5 Association Between Time Spent in Each Intensity Category (SB, LLPA, HLPA, and MVPA) and BMI, Blood Glucose, and HDL Cholesterol for OPACH Study ($N = 4,655$)

	AAI-based intensity				Count-based intensity			
Outcomes	SB	LLPA	HLPA	MVPA	SB	LLPA	HLPA	MVPA
BMI								
Coefficient[a]	0.62	−0.70	−1.52	−1.85	0.65	−0.78	−1.85	−1.22
SE	0.03	0.05	0.06	0.09	0.03	0.05	0.07	0.08
p value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
Glucose								
Coefficient[a]	0.74	−0.66	−1.71	−2.99	0.80	−0.67	−1.82	−2.42
SE	0.14	0.25	0.34	0.50	0.15	0.27	0.39	0.39
p value	<.001	.001	<.001	<.001	<.001	.016	<.001	<.001
HDL cholesterol								
Coefficient[a]	−0.74	1.01	1.46	2.15	−0.79	1.08	1.68	1.61
SE	0.07	0.13	0.18	0.26	0.08	0.14	0.20	0.21
p value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001

Note. Each model was adjusted for average awake wear time (in minutes), age, BMI (except when BMI was the outcome), race/ethnicity, and education. Separate linear regression analyses were conducted based on physical activity metrics derived from AAI and count cutpoints. AAI = accelerometer activity index; OPACH = Objective Physical Activity and Cardiovascular Health; HDL= high-density lipoprotein; SB = sedentary behavior; LLPA = low light physical activity; HLPA = high light physical activity; MVPA = moderate and vigorous physical activity; BMI = body mass index.
[a]Association with 30 min/day incremental difference in physical activity measures.

Table 6 Reallocation of Equivalent Time Spent in Two Activity Intensity Categories in Relation to BMI, Blood Glucose, and HDL Cholesterol for OPACH Study ($N = 4,655$)

	AAI-based intensity			Count-based intensity		
Outcome	SB→LPA	SB→MVPA	LPA→MVPA	SB→LPA	SB→MVPA	LPA→MVPA
BMI						
Coefficient[a]	−0.49	−1.29	−0.81	−0.62	−0.72	−0.10
SE	0.03	0.10	0.11	0.03	0.08	0.10
p value	<.001	<.001	<.001	<.001	<.001	.319
Glucose						
Coefficient[a]	−0.41	−2.59	−2.18	−0.39	−2.16	−1.78
SE	0.17	0.53	0.60	0.19	0.41	0.50
p value	.015	<.001	<.001	.043	<.001	<.001
HDL cholesterol						
Coefficient[a]	0.60	1.56	0.96	0.68	1.15	0.48
SE	0.09	0.27	0.31	0.10	0.21	0.26
p value	<.001	<.001	.002	<.001	<.001	.067

Note. Each model was adjusted for average awake wear time (in minutes), age, BMI (except when BMI was the outcome), race/ethnicity, and education. Separate linear regression analyses were conducted based on physical activity metrics derived from AAI and count cutpoints. AAI = accelerometer activity index; OPACH = Objective Physical Activity and Cardiovascular Health; HDL= high-density lipoprotein; SB = sedentary behavior; LPA = light physical activity; MVPA = moderate and vigorous physical activity; BMI = body mass index.
[a]Effects associated with 30 min/day increase in physical activity measures.

(higher R^2 in calibration equations, e.g., .74 with AAI vs. .54 with counts) and intensity categories (higher AUC in ROC analysis, e.g., 0.92, 0.94, 0.97 with AAI vs. 0.84, 0.88, 0.90 with counts) than counts that were reported in a previous calibration study (Evenson et al., 2015). Both AAI- and counts-based intensity measures can be considered noisy measurements of true activity intensity (e.g., VO_2). For linear regression with imperfect measured predictors, it is well known that regression coefficients are generally attenuated compared with the effects of true exposures and that the magnitude of attenuation bias increases with the magnitude of measurement error (Carroll et al., 2006). Our results suggest that AAI-based measures are associated with substantially less measurement errors than count-based measures, and thus, it is expected that regression based on using AAI-based variables are closer to effects of true activity intensity measures.

This study is limited by the setting, in that it was conducted only in the laboratory which may not fully reflect activities in the real world. Additional free-living collection could help extend the findings. The most significant strength of this study is the calibration of raw accelerometry-based metric that facilitates comparisons between studies collecting raw accelerometry data using different devices. The OPACH study also had a large sample size which enhances measurement precision. In addition, we focused on women 60–91 years old, an understudied age group.

Conclusion

With the availability of raw accelerometry data, nonproprietary metrics have been proposed based on this high-resolution data. Compared with counts data given by the device manufacturer's proprietary algorithms, these raw accelerometry-based metrics make comparisons between studies with different devices possible. However, it is often important to calibrate these new metrics for different populations. In this study, we provide a calibration framework based on AAI, a recently proposed raw accelerometry-based metric, for the OPACH study that included women aged 60–91 years. PA intensity cutpoints were derived from the calibration study within OPACH, and an association analysis between PA intensity and cardiometabolic risk factors demonstrated the effectiveness of the proposed method. Future work could include calibrations with data from broader demographic groups, such as younger age groups and men, and data collected from free-living settings to expand on this work. The results of the present study indicate that AAI is a promising metric for future studies that collect raw accelerometry data.

Acknowledgments

This work was supported by the National Institute of Health, through grants P01AG052352 to AZL (National Institute on Aging, R01HL 130483) to CD (National Heart, Lung, and Blood Institute) and R01CA227122 to KRE (National Cancer Institute, Office of the Director, Office of Disease Prevention, and Office of Behavioral and Social Sciences Research). The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, US Department of Health and Human Services through 75N92021D00001, 75N92021D00002, 75N92021D00003, 75N92021D00004, 75N92021D 00005. All funding agencies were not involved in the design of the study, data collection, analysis, data interpretation, or manuscript writing.

References

2018 Physical Activity Guidelines Advisory Committee. (2018). *2018 Physical activity guidelines advisory committee scientific report*. Department of Health and Human Services. https://health.gov/paguidelines/second-edition/report.aspx

Aittasalo, M., Vaha-Ypya, H., Vasankari, T., Husu, P., Jussila, A.M., & Sievanen, H. (2015). Mean amplitude deviation calculated from raw acceleration data: A novel method for classifying the intensity of adolescents' physical activity irrespective of accelerometer brand. *BMC Sports Science, Medicine and Rehabilitation, 7*(1), 18. https://doi.org/10.1186/s13102-015-0010-0

Anderson, G.L., Manson, J., Wallace, R., Lund, B., Hall, D., Davis S., Shumaker, S., Wang, C.Y., Stein E., & Prentice R.L. (2003). Implementation of the women's health initiative study design. *Annals of Epidemiology, 13*(Suppl. 9), S5–S17. https://doi.org/10.1016/S1047-2797(03)00043-7

Bai, J., Di, C., Xiao, L., Evenson, K.R., LaCroix, A.Z., Crainiceanu, C.M., & Buchner, D.M. (2016). An activity index for raw accelerometry data and its comparison with other activity metrics. *PLoS One, 11*(8), e0160644. https://doi.org/10.1371/journal.pone.0160644

Barone Gibbs, B., Hivert, M.F., Jerome, G.J., Kraus, W.E., Rosenkranz, S.K., Schorr, E.N., Spartano, N.L., Lobelo, F., & American Heart Association Council on Lifestyle and Cardiometabolic Health; Council on Cardiovascular and Stroke Nursing and Council on Clinical Cardiology. (2021). Physical activity as a critical component of first-line treatment for elevated blood pressure or cholesterol: Who, what, and how?: A scientific statement from the American Heart Association. *Hypertension, 78*(2), e26–e37. https://doi.org/10.1161/HYP.0000000000000196

Carroll, R., Ruppert, D., Stefanski, L., & Crainiceanu, C. (2006). *Measurement error in nonlinear models: A modern perspective*. Chapman and Hall.

Chen, K.Y., Janz, K.F., Zhu, W., & Brychta, R.J. (2012). Redefining the roles of sensors in objective physical activity monitoring. *Medicine & Science in Sports & Exercise, 44*(1), S13–S23. https://doi.org/10.1249/MSS.0b013e3182399bc8

Choi, L., Liu, Z., Matthews, C.E., & Buchowski, M.S. (2011). Validation of accelerometer wear and nonwear time classification algorithm. *Medicine & Science in Sports & Exercise, 43*(2), 357–364. https://doi.org/10.1249/MSS.0b013e3181ed61a3

Choi, L., Ward, S.C., Schnelle, J.F., & Buchowski, M.S. (2012). Assessment of wear/nonwear time classification algorithms for triaxial accelerometer. *Medicine & Science in Sports & Exercise, 44*(10), 2009–2016. https://doi.org/10.1249/MSS.0b013e318258cb36

Evenson, K., Wen, F., Herring, A., Di, C., LaMonte, M., Fels Tinker, L., Lee, I.M., Rillamas-Sun, E., LaCroix, A., & Buchner, D. (2015). Calibrating physical activity intensity for hip-worn accelerometry in women age 60 to 91 years: The Women's Health Initiative OPACH Calibration Study. *Preventive Medicine Reports, 2,* 750–756. https://doi.org/10.1016/j.pmedr.2015.08.021

Hastie, T., Tibshirani, R., & Friedman, J. (2017). *The elements of statistical learning: Data mining, inference, and prediction*. Springer.

Jago, R., Zakeri, I., Baranowski, T., & Watso, K. (2007). Decision boundaries and receiver operating characteristic curves: New methods for determining accelerometer cutpoints. *Journal of Sports Sciences, 25*(8), 937–944. https://doi.org/10.1080/02640410600908027

Jakicic, J.M., Powell, K.E., Campbell, W.W., Dipietro, L., Pate, R.R., Pescatello, L.S., Collins, K.A., Bloodgood, B., Piercy, K.L., & 2018 Physical Activity Guidelines Advisory Committee. (2019). Physical activity and the prevention of weight gain in adults: A systematic review. *Medicine & Science in Sports & Exercise, 51*(6), 1262–1269. https://doi.org/10.1249/MSS.0000000000001938

John, D., & Freedson, P. (2012). ActiGraph and Actical physical activity monitors: A peek under the hood. *Medicine & Science in Sports & Exercise, 44*(1), S86–S89. https://doi.org/10.1249/MSS.0b013e3182399f5e

John, D., Tang, Q., Albinali, F., & Intille, S. (2019). An open-source monitor-independent movement summary for accelerometer data processing. *Journal for the Measurement of Physical Behaviour, 2*(4), 268–281. https://doi.org/10.1123/jmpb.2018-0068

Karas, M., Bai, J., Straczkiewicz, M., Harezlak, J., Glynn, N.W., Harris, T., Zipunnikov, V., Crainiceanu, C., & Urbanek, J.K. (2019). Accelerometry data in health research: Challenges and opportunities. *Statistics in Biosciences, 11*(2), 210–237. https://doi.org/10.1007/s12561-018-9227-2

Katzmarzyk, P.T., Powell, K.E., Jakicic, J.M., Troiano, R.P., Piercy, K., Tennant, B., & 2018 Physical Activity Guidelines Advisory Committee. (2019). Sedentary behavior and health: Update from the 2018 physical activity guidelines advisory committee. *Medicine & Science in Sports & Exercise, 51*(6), 1227–1241. https://doi.org/10.1249/MSS.0000000000001935

Kesaniemi, Y., Danforth, E., Jr., Jensen, M., Kopelman, P., Lefebvre, P., & Reeder, B. (2001). Dose-response issues concerning physical activity and health: An evidence-based symposium. *Medicine & Science in*

Sports & Exercise, 33(Suppl. 6), S351–S358. https://doi.org/10.1097/00005768-200106001-00003

Kozey, S., Lyden, K., Staudenmayer, J., & Freedson, P. (2010). Errors in MET estimates of physical activities using 3.5 ml x kg(−1) x min(−1) as the baseline oxygen consumption. *Journal of Physical Activity and Health, 7*(4), 508–516. https://doi.org/10.1123/jpah.7.4.508

LaCroix, A.Z., Rillamas-Sun, E., Buchner, D., Evenson, K.R., Di, C., Lee, I.M., Marshall, S., LaMonte, M.J., Hunt, J., Tinker, L.F., Stefanick, M., Lewis, C.E., Bellettiere, J., & Herring, A.H. (2017). The Objective Physical Activity and Cardiovascular Disease Health in older women (OPACH) study. *BMC Public Health, 17*(1), 192. https://doi.org/10.1186/s12889-017-4065-6

LaMonte, M., Ainsworth, B., & Reis, J. (2006). Measuring physical activity. In T. Wood & W. Zhu (Eds.), *Measurement theory and practice in kinesiology* (pp. 237–271). Human Kinetics.

LaMonte, M.J., Lewis, C.E., Buchner, D.M., Evenson, K.R., Rillamas-Sun, E., Di, C., Lee, I.M., Bellettiere, J., Stefanick, M.L., Eaton, C.B., Howard, B.V., Bird, C., & LaCroix, A.Z. (2017). Both light intensity and moderate-to-vigorous physical activity measured by accelerometry are favorably associated with cardiometabolic risk factors in older women: The Objective Physical Activity and Cardiovascular Health (OPACH) study. *Journal of the American Heart Association: Cardiovascular and Cerebrovascular Disease, 6*(10), e007064. https://doi.org/10.1161/JAHA.117.007064

Liang, K., & Zeger, S. (1986). Longitudinal data analysis using generalized linear models. *Biometrika, 73*(1), 13–22. https://doi.org/10.1093/biomet/73.1.13

Mekary, R.A., Willett, W.C., Hu, F.B., & Ding, E.L. (2009). Isotemporal substitution paradigm for physical activity epidemiology and weight change. *American Journal of Epidemiology, 170*(4), 519–527. https://doi.org/10.1093/aje/kwp163

Ortega, J., & Farley, C. (2007). Individual limb work does not explain the greater metabolic cost of walking in elderly adults. *Journal of Applied Physiology, 102*(6), 2266–2273. https://doi.org/10.1152/japplphysiol.00583.2006

Pepe, M. (2004). *The statistical evaluation of medical tests for classification and prediction*. Oxford University Press.

Reidlinger, D.P., Willis, J.M., & Whelan, K. (2015). Resting metabolic rate and anthropometry in older people: A comparison of measured and calculated values. *Journal of Human Nutrition and Dietetics, 28*(1), 72–84. https://doi.org/10.1111/jhn.12215

Rillamas-Sun, E., Buchner, D.M., Di, C., Evenson, K.R., & LaCroix, A.Z. (2015). Development and application of an automated algorithm to identify a window of consecutive days of accelerometer wear for large-scale studies. *BMC Research Notes, 8*(1), 270. https://doi.org/10.1186/s13104-015-1229-2

Rowlands, A.V. (2018). Moving forward with accelerometer-assessed physical activity: Two strategies to ensure meaningful, interpretable, and comparable measures. *Pediatric Exercise Science, 30*(4), 450–456. https://doi.org/10.1123/pes.2018-0201

Shiroma, E.J., Freedson, P.S., Trost, S.G., & Lee, I.M. (2013). Patterns of accelerometer-assessed sedentary behavior in older women. *The Journal of the American Medical Association, 310*(23), 2562–2563. https://doi.org/10.1001/jama.2013.278896

Stamatakis, E., Rogers, K., Ding, D., Berrigan, D., Chau, J., Hamer, M., & Bauman, A. (2015). All-cause mortality effects of replacing sedentary time with physical activity and sleeping using an isotemporal substitution model: A prospective study of 201,129 mid-aged and older adults. *The International Journal of Behavioral Nutrition and Physical Activity, 12*(1), 121. https://doi.org/10.1186/s12966-015-0280-7

Tzankoff, S.P., & Norris, A.H. (1977). Effect of muscle mass decrease on age-related BMR changes. *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology, 43*(6), 1001–1006. https://doi.org/10.1152/jappl.1977.43.6.1001

van Hees, V.T., Gorzelniak, L., Dean Leon, E.C., Eder, M., Pias, M., Taherian, S., Ekelund, U., Renstrom, F., Franks, P.W., Horsch, A., & Brage, S. (2013). Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity. *PLoS One, 8*(4), e61691. https://doi.org/10.1371/journal.pone.0061691

Yang, C.C., & Hsu, Y.L. (2010). A review of accelerometry-based wearable motion detectors for physical activity monitoring. *Sensors, 10*(8), 7772–7788. https://doi.org/10.3390/s100807772