Calibrating physical activity intensity for hip-worn accelerometry in women age 60 to 91 years: The Women's Health Initiative OPACH Calibration Study

Kelly R. Evenson [a,*], Fang Wen [a], Amy H. Herring [b,c], Chongzhi Di [d], Michael J. LaMonte [e], Lesley Fels Tinker [d], I-Min Lee [f], Eileen Rillamas-Sun [d], Andrea Z. LaCroix [g], David M. Buchner [h]

[a] Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina — Chapel Hill, Chapel Hill, NC, USA
[b] Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina — Chapel Hill, Chapel Hill, NC, USA
[c] Carolina Population Center, University of North Carolina — Chapel Hill, Chapel Hill, NC, USA
[d] Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA
[e] Department of Epidemiology and Environmental Health, School of Public Health and Health Professions, University at Buffalo — SUNY, Buffalo, NY, USA
[f] Division of Preventive Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[g] Division of Epidemiology, Department of Family Medicine and Public Health, University of California, San Diego, La Jolla, CA, USA
[h] University of Illinois at Urbana-Champaign, Champaign, IL, USA

ARTICLE INFO

ABSTRACT

Objective. We conducted a laboratory-based calibration study to determine relevant cutpoints for a hip-worn accelerometer among women ≥60 years, considering both type and filtering of counts.

Methods. Two hundred women wore an ActiGraph GT3X+ accelerometer on their hip while performing eight laboratory-based activities. Oxygen uptake was measured using an Oxycon portable calorimeter. Accelerometer data were analyzed in 15-second epochs for both normal and low frequency extension (LFE) filters. Receiver operating characteristic (ROC) curve analyses were used to calculate cutpoints for sedentary, light (low and high), and moderate to vigorous physical activity (MVPA) using the vertical axis and vector magnitude (VM) counts.

Results. Mean age was 75.5 years (standard deviation 7.7). The Spearman correlation between oxygen uptake and accelerometry ranged from 0.77 to 0.85 for the normal and LFE filters and for both the vertical axis and VM. The area under the ROC curve was generally higher for VM compared to the vertical axis, and higher for cutpoints distinguishing MVPA compared to sedentary and light low activities. The VM better discriminated sedentary from light low activities compared to the vertical axis. The area under the ROC curves were better for the LFE filter compared to the normal filter for the vertical axis counts, but no meaningful differences were found by filter type for VM counts.

Conclusion. The cutpoints derived for this study among women ≥60 years can be applied to ongoing epidemiologic studies to define a range of physical activity intensities.

Introduction

Promoting physically active lifestyles in older adults is recommended for disease prevention and prolonging functionally independent quality life years (Nelson et al., 2007; U.S. Department of Health and Human Services, 2008). Quantifying the amount and intensity of physical activity required for health promotion and disease prevention in older adults is particularly difficult when based on questionnaire assessments (Nelson et al., 2007). Direct measurement of physical activity using accelerometers is gaining interest in public health research and surveillance (Jefferis et al., 2014; Troiano et al., 2008).

In this type of research, the more commonly used accelerometers output the measured accelerations into an integer called "counts" from a single axis (typically the vertical axis) or summarized from multiple axes into a vector magnitude (VM). Counts are averaged over a recording period to summarize movement occurring during that time interval. Among older adults, there currently are no universally accepted cutpoints for defining accelerometer measures into intensity-specific categories of physical activity, which makes interpretation and

comparison of accelerometer measures difficult. Researchers often rely on calibration studies to determine the appropriate count cutpoints or thresholds to define intensity levels of physical activity. Calibration studies help translate counts per unit time into intensity-specific physical activity categories including physical activity intensity categories relevant to specified populations (Jago et al., 2007). However, few studies have focused exclusively on adults 60 years and older, a group for whom intensity-specific count cutpoints may substantially differ from those derived in samples of younger adults, in part due to lower cardiorespiratory fitness, differences in resting metabolic rate, and higher metabolic costs of usual activities of daily living (Evenson et al., 2012).

Because older adults primarily engage in activities of lower intensity (Hooker et al., 2011; Knaggs et al., 2011; Kozey et al., 2010b; Nelson et al., 2007), finer classifications of physical activity at the lower end of the intensity continuum may be useful, particularly because this is reflective of their activities. Furthermore, to better capture light activity and sedentary behavior, the manufacturer of the ActiGraph recently included a low frequency extension (LFE) filter to use during post-processing (ActiGraph, 2015). There is a need to evaluate this in older adults.

While emerging approaches for defining physical activity categories using accelerometry exist, cutpoint-based definitions remain a frequent convention. Thus, we conducted a laboratory-based calibration study to determine cutpoints for a hip-worn accelerometer using both the normal and LFE filters to classify activities as sedentary, light (low and high), and moderate to vigorous physical activity (MVPA) in community-dwelling women ≥60 years.

Methods

Participants

The Objective Physical Activity and Cardiovascular Health (OPACH) Study is an ancillary study of the Women's Health Initiative (WHI) 2010–2015 Long Life Study, which included a calibration substudy to determine intensity-specific accelerometer cutpoints appropriate for this cohort of women. While our sample included only women, to our knowledge no calibration study has indicated cutpoints differ by gender. This substudy was approved by the Institutional Review Boards from each data collection site and by the WHI Clinical Coordinating Center. Participant consent included the possibility to be recruited for the calibration substudy.

In 2013, women from two WHI study centers (Stanford University, University of Alabama — Birmingham) were pre-screened for eligibility to participate in the calibration substudy. Pre-screening exclusion criteria included: (1) a score of ≤1 on the timed walk segment of the Short Physical Performance Battery (Guralnik et al., 1994) indicating a walking limitation; (2) resting systolic blood pressure ≥180 millimeters of mercury (mm Hg); (3) resting diastolic blood pressure ≥110 mm Hg; (4) resting heart rate <40 or >110 beats/min; or (5) use of an assistive device for walking during the home visit.

Eligible study participants were contacted by telephone and further screened. Women who self-reported the following responses were excluded for recruitment: symptoms of chest pain, dizziness, or severe shortness of breath while walking at a usual speed; inability to walk for up to 10 min without using a walker or cane; acute or chronic conditions that would prevent them to walk 400 m; poor balance; and inability to understand questions (suggestive of cognitive impairment). Additionally, in order to meet recruitment goals of 100 women per site, the Birmingham site used word-of-mouth and flyers to recruit community dwelling women not participating in WHI. These non-WHI participants only completed the telephone screen (and not the pre-screening criteria). At the conclusion of the telephone screen, clinic visit appointments were scheduled for all interested participants.

In total, 307 participants were eligible based on OPACH and Long Life Study data for recruitment into the calibration substudy. To recruit 200 participants, the sites called 261 OPACH Study participants from the recruitment list. Reasons for non-participation in the substudy were: ineligible on telephone screen (n = 40), unable to contact (n = 21), not interested (n = 39), canceled (n = 16), or did not complete the study visit (n = 1). Of the 261 participants called, 144 completed the study visit (n = 100 from Stanford where the recruitment list was sufficient for full enrollment into the calibration study and n = 42 from Birmingham where the recruitment list was insufficient for full enrollment into the calibration study). To recruit additional participants, the University of Alabama site screened an additional 62 volunteers until the goal of 58 women to complete the study was reached. Combined, of 323 women contacted, 200 (61.9%) completed the study visit.

Two hundred participants were asked to visit the study clinic site where they signed an informed consent and completed a brief questionnaire. Following this, they performed several standardized physical activities while simultaneously wearing accelerometers, a heart rate monitor, and a portable indirect calorimeter to measure oxygen uptake. Data were collected from both the hip and wrist worn accelerometer; however, this paper focuses solely on data generated from the hip-worn accelerometer.

Measures

All devices used during data collection were synchronized prior to each study visit in order to merge the data properly. The hip-worn accelerometer (ActiGraph GT3X+; Pensacola, Florida) was placed at the iliac crest and secured with a belt. The accelerometer was initialized for use before the first activity task, and data were uploaded from the accelerometer to the computer after the entire visit was completed. The data were output into a file using 15-second epochs (30 Hz) with all three axes for both normal and LFE filters. A 15-second epoch was chosen to attempt to reduce misclassification that may be imposed by longer epoch lengths (Pettee Gabriel et al., 2010a).

VM was derived by taking the square root of the vertical axis squared, plus the anterior–posterior axis squared, plus the medial-lateral axis squared. During data collection, the ActiGraph software (ActiLife) versions 6.4.1, 6.5.2, 6.5.3, and 6.7.1 were used. Oxygen uptake (VO_2) and heart rate were measured continuously during the physical activity tasks using the Oxycon, a portable, battery operated, breath-by-breath metabolic unit (Oxycon Mobile; CareFusion, Rolle, Switzerland). The system consisted of a face mask (Hans Rudolf Inc., Kansas City, Missouri), connected by a one-way non-rebreathing valve and conduction tube to an expired gas analyzer system, and a POLAR heart rate monitor included with the Oxycon. The heart rate monitor was worn on the chest and sent telemetry heart rate data directly to the Oxycon. The oxygen intake system was worn on the back with the weight contributing only minimally to energy expenditure, since it comprised a relatively small percent of total body weight. The Oxycon calibration occurred in the hour before each data collection session with standard gases according to manufacturer specifications. The airflow sensor was calibrated using an automatic 2-point (0.2 L/s and 2 L/s) volume calibration, and gas (O_2/CO_2) analyzers were calibrated using the manufacturer's calibration gas cylinder (5% CO_2, 16% O_2).

Measurements taken at the laboratory visit

Women were asked to not eat or drink caffeinated or calorie containing foods or beverages 2 h prior to the visit, and to not drink alcohol the day of the visit. They were allowed to take medications as usual. To begin, we had participants sit quietly with both feet flat on the floor, without talking for 5 min. After 5 min, we first took a radial pulse for 30 s and then performed the blood pressure measurement. If heart rate was <40 or >110 beats/min, or systolic blood pressure was ≥180 mm Hg, or diastolic blood pressure was ≥110 mm Hg then the

visit ended (which never occurred). Women were asked to report significant health changes since the screening telephone call; three reported health changes but were not sent home because these changes did not prevent participation.

Following a standardized protocol, weight was measured to the nearest 0.5 lb and height was measured to the 0.5 in. using a clinical scale and portable stadiometer. Weight and height were used to calculate body mass index in kilograms per meters squared (kg/m^2). The following categories were used in analyses: underweight <18.5 kg/m^2, normal weight 18.5–<25 kg/m^2, overweight 25–<30 kg/m^2, and obese ≥ 30 kg/m^2 (National Institutes of Health, National Heart Lung and Blood Institute, 1998).

Study activities

We assumed the maximal functional capacity of older women was approximately 6 metabolic equivalents (METs) (Fleg et al., 2005). Therefore, the submaximal physical activity tasks performed in this calibration study were in the range of most women who self-report ability to walk 400 m. With the exception of treadmill walking, participants rested ≥ 2 min between activities so that heart rate could return within 10 beats/min of resting. During this rest period the face mask was removed; when it was time to start the next activity the seal on the mask was rechecked before starting. Simultaneous measurements of accelerometer counts, heart rate, and VO_2 were recorded during the entire period for each physical activity. The duration of activity tasks was chosen to optimize likelihood of achieving steady rate metabolism for measurement of task-specific oxygen uptake. The activities performed were common to women of similar ages (Hooker et al., 2011; Kozey et al., 2010b). The activities were performed in the following order:

– watch DVD while sitting quietly for 7 min,
– wash/dry dishes while standing for 7 min,
– laundry (removing towels from basket and folding) while standing for 7 min,
– 400 meter walk (participant range 1.7 to 10.4 min) (Simonsick et al., 2000; Pettee Gabriel et al., 2010b),
– assemble puzzle while sitting for 7 min,
– dust mopping while standing for 7 min, and
– treadmill walking at two different speeds for 5 min at each speed.

Before any activities, participants were instructed on how to use the Borg rating of perceived exertion (RPE), which was reported just before the end of each activity (Borg and Linderholm, 1974) with numbers ranging from 6 (lowest effort) to 20 (highest effort). Treadmill walking occurred at two different speeds: a slower pace (1.5 miles per hour (mph)) to capture low-intensity walking and a faster pace (2.0 or 2.5 mph) for moderate intensity walking. Determination of a 2.0 mph vs. a 2.5 mph pace for the second walking stage was based on participant's RPE after 5 min into the 1.5 mph walk. Women reporting a RPE of ≤ 11 walked at the 2.5 mph pace, while those reporting a RPE of 12–14 walked at the 2.0 mph pace. Women with a RPE >14 did not continue with the faster pace treadmill walk. Women were asked to not use the treadmill handrails for support during their walk, but were permitted, if needed, to hold the handrails lightly for balance.

Statistical analyses

For each activity performed, descriptive statistics were calculated for measured METs and accelerometer counts. For activities lasting 7 min, we used data observed during minutes 3–7 (4 minute length) for the analysis. For activities lasting 5 min, data from minutes 3–5 were used for the analysis (2 minute length). These time intervals were assumed to align with steady rate oxygen uptake. Among participants who completed the 400-meter walk, we used minute 3 to the end of the walk. Gait speed was calculated by converting the finishing time into m/s.

Measured activity-specific intensity is defined in units of METs, which represent the ratio of activity energy expenditure to resting energy expenditure. The usual definition of resting metabolic rate (1 MET), 3.5 milliliters of oxygen per minute per kilogram ($mL \cdot min^{-1} \cdot kg^{-1}$), is known to be a poor estimate for older adults since it declines with age (Hall et al., 2013; Kozey et al., 2010a). Therefore, our observed activity-specific MET intensity values were calculated using both the usual definition for resting metabolic rate (3.5 $mL \cdot min^{-1} \cdot kg^{-1}$), and another using 3.0 $mL \cdot min^{-1} \cdot kg^{-1}$ to define resting metabolic rate, which was the median value measured in our sample while sitting quietly watching a DVD.

Relationships between accelerometer counts (vertical axis and VM) and measured VO_2, overall and by activities, were assessed graphically and with Spearman correlation coefficients (SCC). Receiver operating characteristic (ROC) curve analysis was conducted to determine vertical axis and VM cutpoints for the intensity categories of sedentary, light (low and high), and MVPA (Jago et al., 2007). For the ROC curve analysis, the breath-by-breath oxygen uptake data were averaged over the time period and, for accelerometry, 15-second data over the time period were used (and not averaged) for the ROC curve analysis. From this, we identified cutpoints using two methods: (i) the maximized sum of sensitivity (maximized by correctly identifying at or above the threshold for intensity) and specificity (maximized by correctly excluding activities below the threshold for intensity) and (ii) the balanced number of false positives and false negatives (identified by calculating the absolute value of the difference between false positives and false negatives and then taking the cutpoint where the minimum absolute value of the difference occurred).

In the ROC curve analysis, the independent variable was the participant's 15-second accelerometer values and the dependent variable was calculated by creating a binary indicator for the calibration activities as either 0 or 1. Four ways to define intensity thresholds were considered in the analysis (Table 1). For example, using the activity classification for sedentary, this corresponded to comparing DVD and putting together a puzzle (assigned a value of 1) versus all other activities (assigned a value of 0). Using the MET value classification for sedentary, this corresponded to counts occurring ≤ 1.5 METs ($=1$) versus >1.5 METs ($=0$). Using the RPE classification for sedentary, this corresponded to counts occurring between 6 and 8 RPE ($=1$) versus ≥ 9 RPE ($=0$). For the MVPA cutpoint, we also explored ROC curve analysis using only the 400-meter walking data because usual walking speed on level ground generally is expected to be of moderate intensity in this age group.

The area under the intensity-specific ROC curve (AUC) was estimated using generalized estimating equations with logistic regression (Liang and Zeger, 1986; Zeger and Liang, 1986), accounting for within-woman correlation due to the same woman participating in multiple activities. This was accomplished using GENMOD with the repeated statement in SAS ® release 9.3 (Cary, North Carolina). The area under the ROC curve represented the accuracy of the test to discriminate between two samples, with values significantly greater than 0.5 indicating better discrimination than by chance alone. General interpretation of the AUC was excellent (0.90–1.00), good (0.80–0.89), fair (0.70–0.79), poor (0.60–0.69), or failure (0.50–0.59). Lastly, leave-one-out cross validation was performed by fitting the models using the cross-validated predicted probabilities from the ROC analysis (Esterman et al., 2010).

Results

Descriptive statistics

Two hundred women completed the study, 100 from each site (142 WHI participants, 58 recruited from the community), with a mean age of 75.5 years (standard deviation 7.7, range 60 to 91 years). Age in decades indicated 21.5% (n = 43) 60–69 years, 44.5% (n = 89) 70–79 years, 32.0% (n = 64) 80–89 years, and 2.0% (n = 4) 90–91 years. Half were non-Hispanic White, while 32.5% were non-Hispanic Black

Table 1
Four ways to define intensity thresholds for the ROC curve analysis.

Intensity	Methods for defining intensity thresholds			
	Based on MET values from each activity (where 1 MET = 3.5 mL/kg/min)	Based on MET values from each activity (where 1 MET = 3.0 mL/kg/min)	Based on Ratings of Perceived Exertion	Based on Specific Activities
Sedentary	≤1.5	≤1.5	6, 7 (easy), or 8	DVD, puzzle
Light low	1.6–2.2	1.6–2.2	9 (very light) or 10	Wash/dry dishes
Light high	2.3–2.9	2.3–2.9	11 (fairly light)	Laundry, mopping, treadmill 1.5 mph[a]
Moderate to vigorous	≥3.0	≥3.0	≥12	400 m walk, treadmill 2.0 mph[a], treadmill 2.5 mph[a]

MET = metabolic equivalent; mL/kg/min: milliliters of oxygen per kilogram per minute; ROC = receiver operating characteristic.

[a] Note: All four methods were calculated with and without treadmill walking. Final results presented in the paper excluded treadmill walking (Tables 3–4). The moderate to vigorous physical activity analyses were also conducted using only the 400 m walk (Table A.3).

and 17.5% were Hispanic. About one-third were normal weight (35.0%), overweight (32.0%), or obese (31.5%), while 3 participants were underweight (1.5%).

When using 3.0 mL·min^{-1}·kg^{-1} to define 1 MET the activities ranged from a mean of 1.0 to 4.0 METs, and when using 3.5 mL·min^{-1}·kg^{-1} to define 1 MET the activities ranged from a mean of 0.8 to 3.4 METs (Table A.1). On average, the 400-meter walk was completed in 6.6 min (standard deviation 1.2), with an average gait speed of 1.0 m/s (standard deviation 0.1). One participant did not attempt the 400-meter walk and 4 stopped early due to fatigue (n = 2), dizziness (n = 1), or for multiple reasons (n = 1). The 400-meter walk yielded the highest mean accelerometer counts among activity tasks (Table 2).

For the treadmill protocol, 174 women walked the first treadmill stage at 1.5 mph (n = 20 did not participate and n = 6 stopped before the 3-minute stage was completed). For the second treadmill stage based on RPE, 68 walked at 2.0 mph (n = 12 stopped before the 3-minute stage was completed), 94 walked at 2.5 mph (n = 3 stopped before the 3-minute stage was completed), and 18 stopped after stage one. Walking on the treadmill at 2.5 mph yielded the highest mean VO$_2$ (12.0 mL·min^{-1}·kg^{-1}).

Accelerometry and METs

The relationship between mean accelerometer counts/15 s and METs by woman for each activity is graphed in Figures A.1 (vertical axis, normal filter), A.2 (vertical axis, LFE filter), A.3 (VM, normal filter), and A.4 (VM, LFE filter). The correlation between measured V.O$_2$ and accelerometer counts was higher (SCC 0.77 to 0.85) than the correlation between RPE and accelerometer counts (SCC 0.50 to 0.62) (Table A.2). Because of concern that the treadmill walking activities may be influenced by extraneous factors such as unfamiliarity, fear of falling, mechanical inefficiency to a greater extent than on-the-ground walking, we repeated the correlation analysis in Table A.2 after removing treadmill walking. This attenuated the correlation for the vertical axis but not for VM (data not shown).

When considering relationships for the groups of activities assigned to an intensity level, the mean MET values (where 1 MET = 3.5 mL·min^{-1}·kg^{-1}) were 1.0 for sedentary (DVD, puzzle), 1.5 for light low (wash/dry dishes), 1.9 for light high (laundry, mopping), and 3.2 for MVPA (400-meter walk); the mean MET values (where 1 MET = 3.0 mL·min^{-1}·kg^{-1}) were 1.2 for sedentary, 1.8 for light low, 2.2 for light high, and 3.8 for MVPA.

ROC curve analysis

Because most women held onto the railing during treadmill walking, the task became unlike usual walking. Our results indicated that the average standard MET value for walking on the treadmill was higher than expected for on-the-ground walking at 1.5 mph. Thus, we excluded treadmill walking from the final ROC curve analysis. Classifying intensity by RPE did not discriminate the intensity categories very well, and so these data are not shown.

The results of the ROC curve analysis are shown in Table 3 (vertical axis) and Table 4 (VM) using three ways to classify intensity (by activity type and by two MET values). For the vertical axis, the two criteria to select cutpoints (maximizing sensitivity/specificity or balancing false positives/false negatives) resulted in generally similar AUC values (0.72–0.79 sedentary, 0.74–0.87 light high, 0.87–0.97 MVPA). The AUC was slightly higher for the LFE filter compared to the normal filter. The cutpoint between sedentary and light low did not discriminate well in

Table 2
Mean and median hip-worn accelerometer results for the vertical axis and vector magnitude by activity; WHI OPACH Calibration Study, 2013.

Physical activity in order of visit	N	Vertical axis in counts/15 s				Vector magnitude in counts/15 s			
		Normal filter		Low frequency extension filter		Normal filter		Low frequency extension filter	
		Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)
Watch DVD	200	0.8 (1.9)	0.0 (0.0, 0.7)	1.3 (2.5)	0.1 (0.0, 1.5)	5.8 (8.2)	2.7 (0.0, 9.2)	7.3 (9.5)	4.1 (0.3, 11.4)
Wash and dry dishes	200	7.2 (12.9)	2.6 (0.2, 8.3)	12.1 (17.5)	5.9 (1.9, 15.0)	60.5 (50.0)	48.6 (25.2, 81.7)	85.7 (58.6)	75.0 (42.2, 112.8)
Laundry while standing	200	21.2 (33.2)	9.1 (3.4, 23.9)	34.2 (42.0)	17.9 (9.0, 41.8)	161.2 (91.5)	135.5 (95.6, 222.0)	206.0 (100.4)	181.3 (132.4, 273.8)
400 meter walk	195	427.3 (178.1)	429.3 (300.7, 540.3)	466.2 (172.0)	464.0 (339.9, 573.8)	654.5 (220.9)	636.8 (518.3, 799.4)	720.5 (215.4)	708.1 (587.5, 851.9)
Assemble puzzle	199	6.1 (14.2)	0.7 (0.0, 3.9)	10.3 (20.1)	2.2 (0.3, 8.6)	30.6 (32.7)	18.8 (9.0, 37.9)	46.3 (40.9)	33.1 (17.5, 57.6)
Mopping	199	82.4 (105.7)	50.5 (19.4, 114.8)	112.0 (111.3)	82.1 (41.9, 154.6)	405.3 (188.6)	365.9 (277.3, 503.4)	465.8 (186.6)	430.0 (343.8, 567.1)
Treadmill (1.5 mph)	174	140.3 (118.7)	111.7 (60.7, 174.6)	190.8 (117.6)	171.1 (115.1, 230.3)	351.0 (203.6)	303.9 (224.7, 419.8)	428.6 (196.6)	382.6 (308.5, 494.1)
Treadmill (2.0 mph)	56	255.2 (166.7)	220.2 (159.1, 311.6)	297.6 (154.0)	264.8 (218.4, 346.7)	450.2 (246.6)	392.2 (299.6, 529.7)	519.3 (235.7)	467.6 (384.4, 591.5)
Treadmill (2.5 mph)	91	416.2 (135.7)	419.0 (315.8, 512.2)	457.7 (129.1)	460.6 (370.3, 550.7)	617.3 (161.8)	613.8 (506.3, 708.0)	686.5 (158.4)	687.2 (577.3, 770.5)

IQR: interquartile range; mph: miles per hour; SD: standard deviation.
Accelerometer data are averaged for each woman and then averaged for the sample.

Table 3
Hip-worn accelerometer cutpoints for the vertical axis (N = 200); WHI OPACH Calibration Study, 2013.

	Sedentary				Light low	Light high				MVPA			
	Se	Sp	AUC	Count/15 s	Count/15 s	Se	Sp	AUC	Count/15 s	Se	Sp	AUC	Count/15 s
Criteria: maximizing the sum of sensitivity plus specificity													
Normal filter													
–Intensity based on activity types	91.4	61.5	0.73	0–0	0–1	76.4	88.1	0.80	1–107	95.1	93.5	0.96	≥108
–Intensity based on measured METS													
Where 1 MET = 3.5 mL/kg/min	84.7	69.0	0.74	0–0	1–30	83.9	85.5	0.86	31–82	94.0	86.0	0.87	≥83
Where 1 MET = 3 mL/kg/min	88.2	61.5	0.72	0–0	0–1	84.1	77.0	0.79	1–81	90.0	90.0	0.90	≥82
Low frequency extension filter													
–Intensity based on activity types	78.8	81.4	0.79	0–0	1–5	85.0	82.1	0.85	6–149	95.4	93.7	0.97	≥150
–Intensity based on measured METS													
Where 1 MET = 3.5 mL/kg/min	78.1	76.6	0.77	0–5	6–48	85.0	84.1	0.87	49–138	90.7	87.8	0.89	≥139
Where 1 MET = 3.0 mL/kg/min	73.3	80.3	0.76	0–0	1–16	82.4	78.2	0.82	17–120	87.6	90.3	0.91	≥121
Criteria: balancing the number of false positives and false negatives													
Normal filter													
–Intensity based on activity types	91.4	61.5	0.73	0–0	0–0	76.4	88.1	0.80	0–193	84.5	97.3	0.96	≥194
–Intensity based on measured METS													
Where 1 MET = 3.5 mL/kg/min	84.7	69.0	0.74	0–0	1–81	74.4	92.8	0.74	82–330	59.9	95.5	0.87	≥331
Where 1 MET = 3.0 mL/kg/min	88.2	61.5	0.72	0–0	1–17	74.4	86.1	0.79	18–190	76.0	95.7	0.90	≥191
Low frequency extension filter													
–Intensity based on activity types	78.8	81.4	0.79	0–0	1–7	82.8	84.3	0.85	8–239	85.2	97.4	0.97	≥240
–Intensity based on measured METS													
Where 1 MET = 3.5 mL/kg/min	76.8	77.9	0.77	0–4	5–11	73.5	92.3	0.87	112–363	58.8	95.3	0.89	≥364
Where 1 MET = 3.0 mL/kg/min	73.3	80.3	0.76	0–0	1–39	73.0	86.3	0.82	40–230	76.0	95.6	0.91	≥231

AUC: area under the curve; MET: metabolic equivalent; mL/kg/min: milliliters of oxygen per kilogram per minute; MVPA: moderate to vigorous physical activity; Se: sensitivity; Sp: specificity.

several cases using normal filter data, indicated by narrow ranges and in some cases values of 0 for both sedentary and light activity.

The VM results indicated generally higher AUC than the vertical axis results (Table 4; 0.84–0.90 sedentary, 0.88–0.94 light high, 0.87–0.92 MVPA). The discrimination between intensity categories was better for VM than vertical axis, and the AUC for VM was similar between the normal and LFE filters. For both the vertical axis and VM results, the LFE filter data resulted in sedentary cutpoints that included higher values and thus a wider range of cutpoints. For both the vertical axis and VM results, the activity-based method for choosing cutpoints generally yielded the highest AUC compared to the other methods of classifying intensity.

For the MVPA cutpoint, we also explored ROC curve analysis using only the 400-meter walking data as this more reflects usual on-the-ground walking (Table A.3). In this case, discrimination was poor (AUC 0.50–0.57 vertical axis, 0.49–0.54 VM). For both the vertical axis and VM, each MVPA cutpoint was identified at a higher accelerometer count threshold.

Cross validation

The leave-one-out cross validation analyses for sedentary, light low, light high, and MVPA showed acceptable results for both the vertical axis and VM, using both the normal and LFE filters (data not shown).

Table 4
Hip-worn accelerometer cutpoints for VM (N = 200); WHI OPACH Calibration Study, 2013.

	Sedentary				Light low	Light high				MVPA			
	Se	Sp	AUC	VM/15 s	VM/15 s	Se	Sp	AUC	VM/15 s	Se	Sp	AUC	VM/15 s
Criteria: maximizing the sum of sensitivity plus specificity													
Normal filter													
–Intensity based on activity types	87.4	79.5	0.88	0–42	43–87	86.4	88.3	0.93	88–305	92.3	84.3	0.92	≥306
–Intensity based on measured METS													
Where 1 MET = 3.5 mL/kg/min	84.2	77.8	0.86	0–87	88–236	89.1	83.4	0.91	237–343	90.2	82.5	0.87	≥344
Where 1 MET = 3.0 mL/kg/min	83.3	76.1	0.84	0–52	53–171	83.1	84.4	0.88	172–295	90.8	83.4	0.90	≥296
Low frequency extension filter													
–Intensity based on activity types	87.1	80.8	0.90	0–65	66–104	89.9	86.0	0.94	105–311	96.4	80.4	0.92	≥312
–Intensity based on measured METS													
Where 1 MET = 3.5 mL/kg/min	81.6	80.5	0.87	0–104	105–299	87.4	84.0	0.90	300–311	93.3	76.0	0.87	≥312
Where 1 MET = 3.0 mL/kg/min	85.6	74.0	0.85	0–95	96–219	82.7	84.1	0.88	220–311	92.5	80.1	0.90	≥312
Criteria: balancing the number of false positives and false negatives													
Normal filter													
–Intensity based on activity types	75.6	87.7	0.88	0–12	13–83	87.0	87.5	0.93	84–521	68.9	94.5	0.92	≥522
–Intensity based on measured METS													
Where 1 MET = 3.5 mL/kg/min	79.8	81.7	0.86	0–62	63–383	79.9	92.3	0.91	384–619	51.6	94.6	0.87	≥620
Where 1 MET = 3.0 mL/kg/min	72.8	84.5	0.84	0–18	19–225	76.7	88.2	0.88	226–518	64.4	93.6	0.90	≥519
Low frequency extension filter													
–Intensity based on activity types	71.2	88.2	0.90	0–31	32–120	87.6	88.0	0.94	121–584	69.6	94.6	0.92	≥585
–Intensity based on measured METS													
Where 1 MET = 3.5 mL/kg/min	79.9	81.9	0.87	0–94	95–439	72.6	92.0	0.90	440–677	50.5	94.3	0.87	≥678
Where 1 MET = 3.0 mL/kg/min	73.2	84.9	0.85	0–39	40–276	76.9	88.2	0.88	277–573	64.5	93.5	0.90	≥574

AUC: area under the curve; MET: metabolic equivalent; mL/kg/min: milliliters of oxygen per kilogram per minute; MVPA: moderate to vigorous physical activity; Se: sensitivity; Sp: specificity; VM: vector magnitude.

Discussion

This study derived hip-worn ActiGraph GT3X+ cutpoints for physical activity intensity categories among older women based on relevant activities performed in the laboratory. Our study contributes by providing intensity-specific cutpoints using the vertical axis or VM, the normal or LFE filters, and a number of sedentary and low intensity activity tasks to enhance characterizing these common activities in older women. This study also contributes to the methodological discussion of how to derive and choose cutpoints using several different criteria. Because it is impractical that a single set of cutpoints could be used in every situation, we presented a range of options for researchers to consider for use.

Four ways the ROC was calculated

This study is potentially the first calibration study to consider four ways to classify the intensity of the activities (Table 1). Most prior calibration studies used either an activity- or MET-based classification, but not both. We also tried using RPE as an additional classifier that captured relative intensity, not usually explored in other studies. Studies indicate reasonable correlation between RPE and measured energy expenditure across a range of work rates, including lower intensity for older adults (Guidetti et al., 2011; Panton et al., 1996). Moreover, RPE is recommended to monitor relative intensity during physical activity of similar energy cost as those in our study (Nelson et al., 2007). Nevertheless, this approach did not perform as well as the activity- or MET-based classification. It may be that there was not enough variability in how women perceived activity intensity to discriminate count thresholds between activities differing in known intensity levels. Using RPE to discriminate intensities may work better in younger adults or when including more vigorous intensity activities that more predictably produce increases in perceived physiological parameters underlying RPE, such as ventilation and local muscle fatigue.

Vertical vs. VM

Accelerometers traditionally output a count only from the vertical plane. Improvements to the ActiGraph were made to provide two other axes, the output from which typically is summarized into a VM. Multi-axial devices are thought to improve measures of complex physical activity patterns that may be incompletely captured by single axis accelerometers. This presumption seems of potentially greater benefit for children, since their movement patterns during free play tend to have shorter sporadic bursts and varied movements in multiple planes (Bailey et al., 1995), but perhaps less important for older adults who may have more constrained and linear movement patterns. Others have indicated that VM can eliminate measurement error due to improper positioning or rotation of the accelerometer. The additional two planes can also help differentiate non-wear from sedentary time, which is important since misclassification of non-wear detrimentally impacts time spent in sedentary behavior (Choi et al., 2012; Hanggi et al., 2013). Our results showed that the use of VM generally resulted in higher AUC's and better discrimination than use of the vertical axis, regardless of the method used to discriminate intensity levels. In several cases, the vertical axis had difficulty discriminating between sedentary and light low intensity levels. This observation suggests that there may be value in using VM over the vertical axis when directly assessing physical activity with a multi-axial accelerometer in older women.

Normal vs. LFE filters

In the study sample, the use of the LFE option did not substantially improve the classification of physical activity intensity categories for VM when evaluated by the AUC. The challenge for the manufacturer is allowing movement through the filter due to human activity and not vibration due to the external environment. Other researchers have found that the LFE filter results in lower non-wearing time and sedentary behavior, but higher levels of mean counts/min, light activity, and MVPA compared to the normal filter (Cain et al., 2013; Wanner et al., 2013). It will be important for future calibration studies to evaluate performance of the LFE filter in other populations with diverse movement tasks (Benka-Wallén et al., 2014).

MVPA cutpoint

The derivation of two MVPA cutpoints (e.g., general lifestyle and walking) is consistent with previous approaches and is based upon research that counts during walking at a given MET level are substantially higher from most other activities at that MET level (Martin et al., 2014; Matthews et al., 2013). The study findings reveal the potential magnitude of differences between MVPA cutpoints derived using lifestyle activities and MVPA cutpoints using only the 400-meter walk. The calibration study activities proposed were not expected to exceed the conventional criteria for defining vigorous intensity on an absolute scale (e.g., 6 METs). Other calibration studies of older adults have been conducted similarly (Copeland and Esliger, 2009; Hooker et al., 2011). Thus, we only determined a threshold for activities of at least moderate intensity rather than activities pre-specified as being both moderate and vigorous intensity (e.g., MVPA and not separately for vigorous activity). We found the types of activities impacted the derived cutpoints, with activities focusing on ambulation providing higher MVPA thresholds than those observed for the other lifestyle activities. Both cutpoints may be useful, however, and their application may be specific to the circumstance in which physical activity is being assessed. The values used are also impacted by which method is applied to select cutpoints.

Choosing cutpoints

We derived cutpoints by methods that either (i) maximized the sum of sensitivity and specificity or (ii) balanced the number of false positives and false negatives. The choice between these two methods substantially impacted the specific cutpoints for light high and MVPA. In prior studies, the more common approach was to choose cutpoints by only maximizing the sum of sensitivity and specificity, or equivalently minimizing the overall misclassification rate (Evenson et al., 2008). However, this approach might suffer from low positive predicted values, especially for MVPA where we had fewer observations. The second approach of balancing false positives and false negatives is preferable for the purpose of predicting summary variables, such as light low, light high, or MVPA minutes. The cutpoints derived from this approach provide roughly unbiased estimates for these summary variables. Other criteria may be specified to generate alternative cutpoints.

Limitations and strengths

This study had several limitations. First, activities were performed in a laboratory setting rather than in a real-world setting. We did encourage women to perform activities as they usually would, but also within the protocol description of the activity in order to be standardized across women. Second, we did not obtain direct measures of peak functional capacity in order to evaluate accelerometer cutpoints relative to physical fitness level. Because of the known decline in maximal VO_2 with aging (Fleg et al., 2005), use of absolute intensity cutpoints (e.g., MET values) may lead to meaningful misclassification of activity levels in older adults. Third, we did not obtain a measure of resting metabolic rate. The strengths of this study included the large sample size, a focus on women age 60 to 91 years, and the variety of ways in which the data were explored (by filter, by vertical axis and VM, by four methods for defining intensity thresholds). We included a variety of activity intensities and tasks, used a tri-axial accelerometer, and provided cutpoints for both the vertical axis and VM.

Conclusion

While individualized thresholds may be preferred for older adults (Pruitt et al., 2008), it is often not feasible in public health studies nor for providing clinical guidance. As a fast growing segment of the US population at risk of several chronic diseases by virtue of age alone, research into physical activity is both timely and critical. This study provides cutpoints for women 60–91 years to distinguish varying levels of intensity while wearing the hip-worn ActiGraph accelerometer. Our study explored thresholds at two MET levels of resting energy expenditure and two accelerometer filter sensitivities for both the vertical axis and VM. Future work could determine how to better use the raw signal using data collected multiple times per second to determine not only intensity but also position and pattern of physical activity (Lyden et al., 2011; Staudenmayer et al., 2009). Future work could also include activities performed outside of the laboratory setting and activities of higher intensity.

Conflicts of interest statement

The authors declare that there are no conflicts of interest.

Sources of funding

Acknowledgment

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.pmedr.2015.08.021.

References

ActiGraph, 2015. Low frequency extension explained Accessed July 28, 2015 at https://help.theactigraph.com/entries/21767838-Low-Frequency-Extension-Explained.

Bailey, R., Olson, J., Pepper, S., Porszasz, J., Barstow, T., Cooper, D., 1995. The level and tempo of children's physical activities: an observational study. Med. Sci. Sports Exerc. 27, 1033–1041.

Benka-Wallén, M., Nero, H., Franzen, E., Hagstromer, M., 2014. Comparison of two accelerometer filter settings in individuals with Parkinson's disease. Physiol. Meas. 35, 2287. http://dx.doi.org/10.1088/0967-3334/35/11/2287.

Borg, G., Linderholm, H., 1974. Perceived exertion and pulse rate during graded exercise in various age groups. Acta Med. Scand. 472, 194–206.

Cain, K.L., Conway, T.L., Adams, M.A., Husak, L.E., Sallis, J.F., 2013. Comparison of older and newer generations of ActiGraph accelerometers with the normal filter and the low frequency extension. Int. J. Behav. Nutr. Phys. Act. 10, 51 (Available at http://www.ijbnpa.org/content/10/1/51).

Choi, L., Ward, S.C., Schnelle, J.F., Buchowski, M.S., 2012. Assessment of wear/nonwear time classification algorithms for triaxial accelerometer. Med. Sci. Sports Exerc. 44, 2009–2016.

Copeland, J., Esliger, D., 2009. Accelerometer assessment of physical activity in active, healthy older adults. J. Phys. Act. Health 17, 17–30.

Esterman, M., Tamber-Rosenau, B.J., Chiu, Y.C., Yantis, S., 2010. Avoiding non-independence in fMRI data analysis: leave one subject out. NeuroImage 50, 572–576.

Evenson, K.R., Catellier, D., Gill, K., Ondrak, K., McMurray, R., 2008. Calibration of two objective measures of physical activity for children. J. Sport Sci. 26, 1557–1565.

Evenson, K., Buchner, D., Morland, K., 2012. Objective measurement of physical activity and sedentary behavior among United States adults 60 years and older. Prev. Chronic Dis. 9, 1–10 (Available at http://www.cdc.gov/pcd/issues/2012/2011_0109.htm).

Fleg, J.L., Morrell, C.H., Bos, A.G., et al., 2005. Accelerated longitudinal decline of aerobic capacity in healthy older adults. Circulation 112, 674–682.

Guidetti, L., Sgadari, A., Buzzachera, C.F., et al., 2011. Validation of the OMNI-cycle scale of perceived exertion in the elderly. J. Aging Phys. Act. 19, 214–224.

Guralnik, J., Simonsick, E., Ferrucci, L., et al., 1994. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J. Gerontol. 49, M85–M94.

Hall, K.S., Howe, C.A., Rana, S.R., Martin, C.L., Morey, M.C., 2013. METs and accelerometry of walking in older adults: standard versus measured energy cost. Med. Sci. Sports Exerc. 45, 574–582.

Hanggi, J.M., Phillips, L.R., Rowlands, A.V., 2013. Validation of the GT3X ActiGraph in children and comparison with the GT1M ActiGraph. J. Sci. Med. Sport 16, 40–44.

Hooker, S.P., Feeney, A., Hutto, B., et al., 2011. Validation of the actical activity monitor in middle-aged and older adults. J. Phys. Act. Health 8, 372–381.

Jago, R., Zakeri, I., Baranowski, T., Watson, K., 2007. Decision boundaries and receiver operating characteristic curves: new methods for determining accelerometer cutpoints. J. Sports Sci. 25, 937–944.

Jefferis, B.J., Sartini, C., Lee, I.M., et al., 2014. Adherence to physical activity guidelines in older adults, using objectively measured physical activity in a population-based study. BMC Public Health 14, 382 (Available at http://www.biomedcentral.com/1471-2458/14/382).

Knaggs, J.D., Larkin, K.A., Manini, T.M., 2011. Metabolic cost of daily activities and effect of mobility impairment in older adults. J. Am. Geriatr. Soc. 59, 2118–2123.

Kozey, S., Lyden, K., Staudenmayer, J., Freedson, P., 2010a. Errors in MET estimates of physical activities using $3.5 \ ml \times kg^{-1} \times min^{-1}$ as the baseline oxygen consumption. J. Phys. Act. Health 7, 508–516.

Kozey, S.L., Lyden, K., Howe, C.A., Staudenmayer, J.W., Freedson, P.S., 2010b. Accelerometer output and MET values of common physical activities. Med. Sci. Sports Exerc. 42, 1776–1784.

Liang, K., Zeger, S., 1986. Longitudinal data analysis using generalized linear models. Biometrika 73, 13–22.

Lyden, K., Kozey, S., Staudenmeyer, J.W., Freedson, P.S., 2011. A comprehensive evaluation of commonly used accelerometer energy expenditure and MET prediction equations. Eur. J. Appl. Physiol. 111, 187–201.

Martin, K.R., Koster, A., Murphy, R.A., et al., 2014. Changes in daily activity patterns with age in U.S. men and women: National Health and Nutrition Examination Survey 2003–04 and 2005–06. J. Am. Geriatr. Soc. 62, 1263–1271.

Matthews, C.E., Keadle, S.K., Sampson, J., et al., 2013. Validation of a previous-day recall measure of active and sedentary behaviors. Med. Sci. Sports Exerc. 45, 1629–1638.

National Institutes of Health, National Heart Lung and Blood Institute, 1998. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: the evidence report. Obes. Res. 6, S51–S209.

Nelson, M., Rejeski, W., Blair, S., et al., 2007. Physical activity and public health in older adults. Recommendation from the American College of Sports Medicine and the American Heart Association. Circulation 116, 1094–1105.

Panton, L.B., Graves, J.E., Pollock, M.L., et al., 1996. Relative heart rate, heart rate reserve, and VO_2 during submaximal exercise in the elderly. J. Gerontol. Med. Sci. 51A, M165–M171.

Pettee Gabriel, K.K., McClain, J.J., Schmid, K.K., et al., 2010a. Issues in accelerometer methodology: the role of epoch length on estimates of physical activity and relationships with health outcomes in overweight, post menopausal women. Int. J. Behav. Nutr. Phys. Act. 7, 53 (Available at http://www.ijbnpa.org/content/7/1/53).

Pettee Gabriel, K.K., Rankin, R.L., Lee, C., Charlton, M.E., Swan, P.D., Ainsworth, B.E., 2010b. Test–retest reliability and validity of the 400-meter walk test in healthy, middle-aged women. J. Phys. Act. Health 7, 649–657.

Pruitt, L.A., Glynn, N.W., King, A.C., et al., 2008. Use of accelerometry to measure physical activity in older adults at risk for mobility disability. J. Aging Phys. Act. 16, 416–434.

Simonsick, E.M., Gardner, A.W., Poehlman, E.T., 2000. Assessment of physical function and exercise tolerance in older adults: reproducibility and comparability of five measures. Aging 12, 274–280.

Staudenmayer, J., Pober, D., Crouter, S., Bassett, D., Freedson, P., 2009. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. J. Appl. Physiol. 107, 1300–1307.

Troiano, R., Berrigan, D., Dodd, K., Masse, L., Tilert, T., McDowell, M., 2008. Physical activity in the United States measured by accelerometer. Med. Sci. Sports Exerc. 40, 181–188.

U.S. Department of Health and Human Services, 2008. 2008 Physical Activity Guidelines for Americans. ODPHP Publication No. U0036. Washington, D.C.pp. 1–61 (Available at http://www.health.gov/paguidelines/).

Wanner, M., Martin, B.W., Meier, F., Probst-Hensch, N., Kriemler, S., 2013. Effects of filter choice in GT3X accelerometer assessments of free-living activity. Med. Sci. Sports Exerc. 45, 170–177.

Zeger, S., Liang, K., 1986. Longitudinal data analysis for discrete and continuous outcomes. Biometrics 42, 121–130.