Received: 15 September 2022 | Revised: 3 May 2023 | Accepted: 19 May 2023

DOI: 10.1111/sms.14416

ORIGINAL ARTICLE

WILEY

Absolute intensity thresholds for tri-axial wrist and waist accelerometer-measured movement behaviors in adults

Gregore Iven Mielke[1] | **Márcio de Almeida Mendes**[2] | **Ulf Ekelund**[3,4] | **Alex V. Rowlands**[5,6] | **Felipe Fossati Reichert**[7] | **Inacio Crochemore-Silva**[2,7]

[1]School of Public Health, The University of Queensland, Queensland, Brisbane, Australia

[2]Post-graduate Program in Epidemiology, Federal University of Pelotas, Pelotas, Brazil

[3]Norwegian School of Sport Sciences, Oslo, Norway

[4]Department of Chronic Diseases, Norwegian Institute of Public Health, Oslo, Norway

[5]Diabetes Research Centre, University of Leicester, Leicester General Hospital, Leicester, UK

[6]NIHR Leicester Biomedical Research Centre, Leicester, UK

[7]Post-graduate Program in Physical Education, Federal University of Pelotas, Pelotas, Brazil

Correspondence
Gregore Iven Mielke, School of Public Health, The University of Queensland | Herston Campus, Brisbane, QLD 4006, Australia.
Email: g.ivenmielke@uq.edu.au

Funding information
Conselho Nacional de Desenvolvimento Científico e Tecnológico; National Health and Medical Research Council; National Institute for Health Research

Aim: This study was aimed to: (1) compare raw triaxial acceleration data from GENEActiv (GA) and ActiGraph GT3X+ (AG) placed on the non-dominant wrist; (2) compare AG placed on the non-dominant and dominant wrist, and waist; (3) derive brand- and placement-specific absolute intensity thresholds for inactive and sedentary time, and physical activity intensity in adults.

Methods: Eighty-six adults (44 men; 34.6 ± 10.8 years) performed nine activities while simultaneously wearing GA and AG on wrist and waist. Acceleration (in gravitational equivalent units; mg) was compared with oxygen uptake (measured with indirect calorimetry).

Results: Increases in acceleration mirrored increases in intensity of activities, regardless of device brand and placement. Differences in acceleration between GA and AG worn at the non-dominant wrist were small but tended to be high at lower intensity activities. Thresholds for differentiating inactivity (<1.5 MET) from activity (≥1.5 MET) ranged from 25 mg (AG non-dominant wrist; sensitivity 93%, specificity 95%) to 40 mg (AG waist; sensitivity 78%, specificity 100%). For moderate intensity (≥3 METs), thresholds ranged from 65 mg (AG waist; sensitivity 96%, specificity 94%) to 92 mg (GA non-dominant; sensitivity 93%, specificity 98%); vigorous intensity (≥6 METs) thresholds ranged from 190 mg (AG waist; sensitivity 82%, specificity 92%) to 283 mg (GA non-dominant; sensitivity 93%, specificity 98%).

Conclusion: Raw triaxial acceleration outputs from two widely used accelerometer brands may have limited comparability in low intensity activities. Thresholds derived in this study can be utilized in adults to reasonably classify movement behaviors into categories of intensity.

KEYWORDS

accelerometer, adults, calibration, physical activity, sedentary, validity

1 | INTRODUCTION

Accurate measurement of movement behaviors in population-based studies is crucial to ensure high-quality evidence on prevalence and trends, correlates and determinants, and health consequences of physical activity and sedentary time.[1] Despite device-measured sedentary time and physical activity is often referred to as an 'objective' measure,[1] physical activity and sedentary time derived from accelerometers can be largely influenced by methodological decisions, hence limiting the comparability of these estimates within and between studies.[2,3] For example, device brand, wear-site, and definition of intensity thresholds have been shown to affect individual and population estimates of time spent in sedentary behavior and different intensities of physical activity in epidemiological studies.[2]

Development of open-access codes such as GGIR[4] for processing, analysis and conversion of raw accelerometry data into estimates of time spent in different intensities of movement behaviors has facilitated the comparability between studies and become an alternative to data processing using count-based brand-specific cut-points.[4] Moreover, the availability of raw data has improved comparability between studies collecting physical activity using different device brands such as GENEActiv (GA) and ActiGraph (AG), which have been used widely in large population-based studies.[2,3,5,6]

Although GA and AG provide highly correlated raw acceleration data,[6] findings from previous studies conducted in the UK and Norway have suggested that outcomes based on raw acceleration provided by these device brands may not be entirely comparable.[6–9] GA accelerometers appeared to provide higher acceleration values than AG, especially when assessing activities with low acceleration outputs.[6–9] These differences may limit the use of thresholds that are not device appropriated for classification of physical activity intensities. Moreover, given that acceleration estimates produced in different continents may be subject to calibration error as a potential source of bias, despite its small magnitude,[10] calibration studies to compare acceleration outputs between device-brands and derived intensity thresholds are particularly important in locations other than Europe. More research is needed to advance knowledge about the similarities and differences of acceleration outputs at different activities between these devices.

The pioneer study to derive sedentary time and physical activity intensity thresholds for raw accelerometer data was conducted by Hildebrand and colleagues with a sample of 30 young adults and 30 children in Norway.[9] In their study, raw triaxial accelerometer outputs were compared between GA and AG worn at the waist and non-dominant wrist during eight different activities and used to develop regression equations for estimating energy expenditure from raw accelerometer output using indirect calorimetry as the reference method. The thresholds developed by Hildebrand and colleagues have been widely used to classify sedentary time[11] and moderate-to-vigorous physical activities[9] (>720 citations; Google Scholar, March 2023), including large population-based studies such as the UK Biobank[12] and NHANES[13] in high-income countries, and the Pelotas (Brazil) Birth Cohorts as an example of low/middle-income setting.[14–16]

Comparing derived thresholds for inactive/sedentary time and physical activity intensities between studies may be challenging because even small changes in definitions (e.g., sedentary time defined by energy expenditure and posture, or inactive time defined by energy expenditure only) and in the analytical approach can affect the accuracy of discrimination between intensities. Given the importance of replication in science, as well as the necessary evidence of comparison between accelerometer-brands and physical activity behaviors thresholds from different continents, we conducted a calibration study protocol based on Hildebrand et al in Pelotas (Brazil). Therefore, the aims of this study were to: (1) compare raw triaxial acceleration from GENEActiv (GA) and ActiGraph GT3X+ (AG) placed on the non-dominant wrist; (2) compare raw triaxial acceleration from AG placed on the non-dominant and dominant wrist, and waist; and (3) derive brand- and placement-specific absolute intensity thresholds for inactive and sedentary time, and physical activity intensity in adults.

2 | METHODS

2.1 | Sample

This study included data of 86 adults from Pelotas, Brazil, recruited through various social outlets for the purpose of the study. Multiple strategies were used to recruit participants, including distribution of flyers, emails, and word of mouth. To improve external validity of the findings, the recruitment of participants was targeted to include a balanced sample in terms of gender (44 men; 42 women), age (mean age: 34.6 years; standard deviation: 10.8; age range: 20–59 years) and participation in leisure-time physical activity or active transportation, which was used for recruitment purposes as a proxy of fitness level. Participation in leisure-time physical activity and active transportation was obtained during the screening interview, by asking participants whether they regularly engaged in physical activities during their leisure time or for commuting purposes. Based on their responses, they were classified as either active or non-active. Individuals with long term health conditions, including diabetes, cardiovascular or pulmonary diseases were not included in the study. All participants had no contraindications to participation in physical activity and were able to perform daily living activities. Written informed consent was obtained from all participants. This study received approval from the Ethics Committee of the Scholl of Medicine from the Federal University of Pelotas (Protocol #1.258.787/2015). Further details of the study protocol have been published elsewhere.[17]

2.2 | Protocol of activities and measurement

Participants were invited to attend to laboratory-based protocol in the Laboratory of Exercise Physiology in the School of Physical Education at the Federal University of Pelotas (Brazil). Participants undertook a protocol that included nine activities common in free-living conditions. The session started with participants *lying* in supine position, with the arms at the side for 10 min. Afterwards, participants performed eight activities for 5 min each in the following sequence: (1) *sitting* in a chair using a computer; (2) *standing* on the floor using mobile phone; (3) *circuit*, which included sitting, putting on shoes, standing, moving eight light objects on a desk, texting on a mobile phone and sitting down; (4) *slow walking* at 3 km/h; (5) *brisk walking* at 6 km/h; (6) *walking up a step* (20 cm) 15 times followed by 1-min brisk walking in a treadmill; (7) *running* at 8 km/h; and (8) *intermittent running*, which included running at 10 km/h for 1 min followed by running at 12 km/h for 30 s until the completion. Two minutes of resting between the activities were allowed, with exception of running and intermittent running, for which a break of 5 min was allowed. With exception of intermittent running, this protocol used activities that were similar to those used in the seminal study by Hildebrand et al.[9]

Participants had their oxygen uptake continuously measured using a portable gas analyzer (VO2000, MedGraphics; Ann Arbor, USA) that was previously calibrated according to the manufacturer's instructions to ensure accuracy. Oxygen uptake was measured once every three breaths, and the data were analyzed using BREEZE software. The average of the oxygen uptake ($ml \cdot kg^{-1} \cdot min^{-1}$) of each activity was calculated and converted to metabolic equivalent of task (MET; 1 MET = 3.5 $mL \cdot kg^{-1} \cdot min^{-1}$). Additionally, the $\dot{V}O_2$ data converted to METs were coded into four absolute-intensity categories widely described in the literature[18]: inactive (<1.5 MET), light (1.5–2.9 METs), moderate (3.0–5.9 METs); vigorous (≥6.0 METs). Participants wore four accelerometers while the testing session was performed. These included ActiGraph wGT3X-BT model (ActiGraph LLC) (AG) accelerometers placed on the waist, dominant and non-dominant wrist, and one GENEActiv (GA) on the non-dominant wrist. Due to the availability of devices, GA was only placed on the non-dominant wrist. The order of placement, in which the two devices were first placed at the distal side of the non-dominant wrist joint, was randomized. Both devices measured body acceleration in three axes at 60 Hz, according to the manufacturer's descriptions. To ensure the inclusion of activities in steady state, only the average oxygen uptake, as well as the average acceleration in milligravitational units (mg) of the last 2 min of each activity were considered in the analyses.

2.3 | Accelerometer data reduction

Accelerometers were set up and data downloaded using their correspondent commercially available software (ActiLife software version 6.5.1, USA, and GENEActiv personal computer software version 2.2, UK). Raw data were processed in R using GGIR package version 1.7–1.[4] Data were initially aggregated in 5-s time series and the vector magnitude of the three axes was used to calculate activity-related acceleration in mg using Euclidian Norm minus 1 g [$ENMO = \sqrt{(x^2 + y^2 + z^2)}$-1].[10]

2.4 | Statistical analysis

Initially, means, medians, standard deviations, and interquartile ranges of MET values and acceleration were calculated for each activity performed. Box and whisker plot graphs were used to visually present the distribution of acceleration outputs during each activity performed. To address the aims of the study, statistical analyses were conducted in multiple steps. Different analytical approaches were used to identify and validate thresholds for sedentary time and inactive, and physical activities of different intensities.

2.4.1 | Comparison between accelerometry outputs from GENEActiv and ActiGraph and device placement

Bland–Altman methods[19,20] were used to examine the agreement between outputs from different brands (GA and AG), worn at the same wear-site (non-dominant wrist), and between outputs from the same brand (AG) worn at three wear-sites (waist, dominant wrist, and non-dominant wrist). In addition, Lin's Concordance

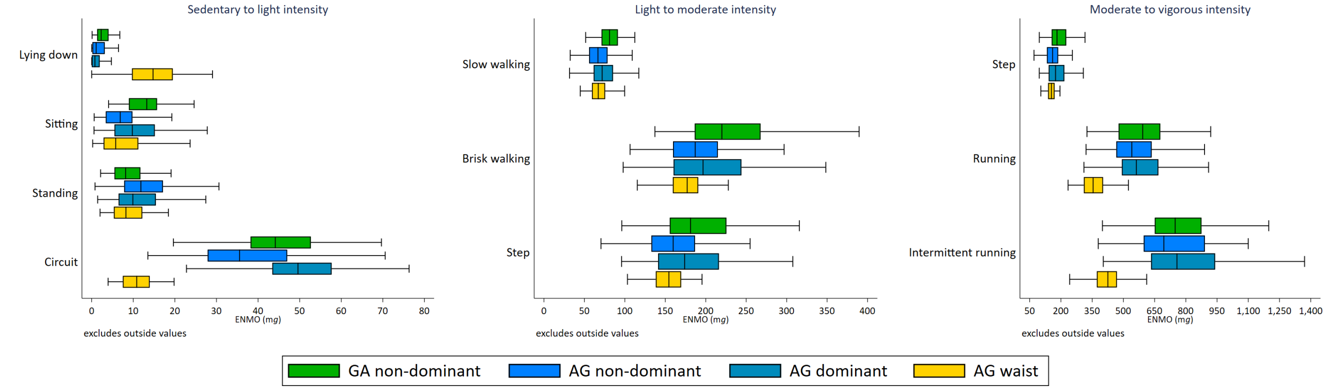

FIGURE 1 Box and whiskers plots of the distribution of acceleration outputs from GENEActiv (GA) and ActiGraph (AG) during each activity performed. Due to the wide range in acceleration, and to enhance visualization, figures were rendered using different scales on the X-axis. For visual and comparative purposes, 'Step' was included twice in the plots (light to moderate intensity; moderate to vigorous intensity). Green and bright blue plots represent brand comparison; bright blue, dark blue and yellow plots represent wear-site comparison for AG. Outside (25th centile − 1.5 × range; 75th centile + 1.5 interquartile range) values were excluded from the figure for visual purposes.

Correlation Coefficient (CCC),[21] mean differences in acceleration and respective 95% limits of agreement between outputs were calculated for each activity, and according to four categories of intensity based on MET values: (a) <1.5 MET; (b) 1.5–2.9 METs; (c) 3.0–5.9 METs; (d) ≥6.0 METs. Bland–Altman plots were used and regression-based limits of agreement for differences in acceleration outputs were used to determine whether agreement between outputs varied by level of acceleration.[19]

2.4.2 | Thresholds for inactive and sedentary time

Absolute thresholds for inactive time and sedentary time were estimated using different approaches. Inactive time was defined based on energy expenditure (<1.5 MET), whereas sedentary time was defined based on energy expenditure (<1.5 MET) and posture (lying or sitting).[22] For inactive time, activities were recoded based on MET values to create a binary variable used to differentiate inactive (<1.5 MET) from active time (≥1.5 METs). For sedentary time, analyses were conducted to add comparability between findings of this study with the previous study by Hildebrand.[11] Thresholds for sedentary time were estimated by comparing acceleration outputs generated during lying and sitting with acceleration outputs of standing, slow walking, brisk walking and running. Receiver operating characteristics (ROC) curves were used to identify acceleration thresholds able to differentiate inactive from active time, and sedentary from non-sedentary time. ROC area, sensitivity and specificity were calculated for each level of raw acceleration and for each device/placement. The optimal thresholds for inactive and sedentary time were defined by using the Youden Index, which was calculated as the sum of sensitivity and specificity.

2.4.3 | Thresholds for moderate and vigorous intensity activities

Initially, linear regression models were fitted to examine the relationship of acceleration outputs from each device and placement (independent variable) with MET values (dependent variable). In addition, a quadratic polynomial term for acceleration was included in the regression models to account for the non-linear relationship observed between acceleration outputs and MET values. Predictive equations for estimating MET values based on acceleration outputs were generated, and thresholds for moderate (3 MET) and vigorous (6 MET) intensity physical activity were estimated from the regression equations. To allow comparability between the findings of this study with findings of the seminal study by Hildebrand,[9] regression models were also fitted without including the quadratic polynomial term and intermittent running in the analyses (Appendix S1).

2.4.4 | Accuracy of optimal intensity thresholds

Optimal intensity thresholds identified in this study were applied to the sample to assess the intensity classification accuracy for intensity. Thus, based on accelerometry outputs, activities were categorized into one of the four categories of absolute intensity (inactive or sedentary, light, moderate, vigorous) using the optimal thresholds identified for each accelerometer brand/placement. Sensitivity, specificity, and area under the ROC curve were calculated for each intensity. Additionally, an average confusion matrix and kappa statistics were calculated to indicate the ability of thresholds to accurately classify activities (% accurately classified). All statistical analyses were performed using STATA 17.1.

3 | RESULTS

Descriptive characteristics of the sample are presented in Table S1. The mean (SD) height and weight of the participants were 169.4 (9.2) cm and 72.5 (14.0) kg, respectively. The distribution of acceleration outputs during each activity is presented in Figure 1, while the distribution of oxygen consumption and MET values during each activity are presented in Table 1. Overall, increases in acceleration mirrored increases in intensity of activities, regardless of placement and brand of the accelerometer. However, changes in the average and variability of acceleration varied by placement. Distribution of acceleration outputs from waist placement was more homogenous than that from wrist accelerometers, except for lying down and to a lesser extent sitting down. In addition, comparing only dominant with non-dominant wrist, except for circuit activity, the non-dominant wrist had a more homogeneous distribution than dominant-wrist assessment. Acceleration outputs and the shape of the distribution for sitting were similar to standing, both within and between accelerometers. For circuit activities, running and intermittent running, differences in the average and distribution of acceleration between placements were observed. Overall, the average acceleration from the waist accelerometer was lower than from wrist accelerometers for moderate to vigorous intensity activities.

3.1 | Comparison between accelerometry outputs from GENEActiv and ActiGraph and device placement

Bland–Altman plots of the agreement between acceleration outputs (mg) according to MET values are presented in Figure 2 and Table S2. The mean difference in non-dominant wrist accelerometer data between outputs from GA and AG was 21 mg (higher in GA than AG; the top left panel of Figure 2). The concordance correlation coefficient was 0.98 (95% CI: 0.97–0.99). Overall, the average mean

TABLE 1 Physiological parameters and accelerometer outputs from GENEActiv (GA) and ActiGraph (AG) during each activity performed.

	N[a]	METs	VO_2[b]	Mean acceleration (SD)			
		Mean (SD)	Mean (SD)	GA non-dominant	AG non-dominant	AG dominant	AG waist
Activities							
Lying down[c]	70	1.0 (0.2)	3.5 (0.7)	2.7 (2.0)	2.8 (4.7)	1.6 (2.6)	15.2 (7.1)
Sitting	78	1.3 (0.2)	4.5 (0.9)	13.6 (6.6)	7.5 (5.0)	10.9 (6.5)	8.4 (8.3)
Standing	76	1.2 (0.3)	4.4 (1.0)	10.1 (7.3)	13.2 (8.0)	11.9 (7.9)	9.0 (4.9)
Circuit	77	2.0 (0.4)	7.0 (1.5)	45.7 (11.5)	38.2 (12.7)	49.7 (13.0)	11.2 (4.5)
Slow walking	76	3.1 (0.6)	10.7 (2.0)	84.2 (24.1)	68.7 (19.6)	74.6 (20.9)	69.1 (11.9)
Brisk walking	76	5.4 (1.0)	19.0 (3.4)	234.6 (73.3)	193.4 (53.8)	213.0 (70.1)	179.5 (25.2)
Step	75	5.5 (0.8)	19.1 (2.9)	201.7 (68.6)	169.7 (55.3)	187.8 (59.9)	157.2 (23.6)
Running	71	8.2 (1.1)	28.8 (3.8)	602.9 (179.0)	575.4 (157.4)	605.7 (152.9)	365.6 (66.0)
Intermittent running	46	10.5 (1.4)	36.7 (4.9)	765.2 (216.9)	729.5 (201.8)	804.5 (225.8)	425.3 (82.3)
Intensities based on MET values							
<1.5 METs	206	1.1 (0.2)	4.0 (0.7)	9.4 (8.9)	8.7 (9.0)	8.9 (10.3)	11.2 (7.5)
1.5–2.9 METs	136	2.2 (0.5)	7.6 (1.6)	52.7 (24.8)	41.5 (23.1)	50.8 (24.6)	28.1 (28.3)
3.0–5.9 METs	159	4.8 (0.9)	16.5 (3.0)	193.6 (96.9)	169.9 (116.6)	183.4 (117.4)	152.2 (65.0)
≥6 METs	144	8.7 (1.6)	30.5 (6.0)	580.9 (261.4)	539.9 (250.2)	581.5 (270.7)	340.9 (114.2)

Abbreviations: AG, ActiGraph; GA: GENEActiv; SD, standard deviation.

[a]Number of observations. For the activities listed, each observation refers to a single participant engaging in that activity. For intensities categorized based on MET values, the number of observations reflects the number of participant-activity pairs that fall within a particular MET range.

[b]VO_2 is expressed in milliliters per kilogram per minute (mL O^2/kg/min).

[c]Activity performed for 10 min.

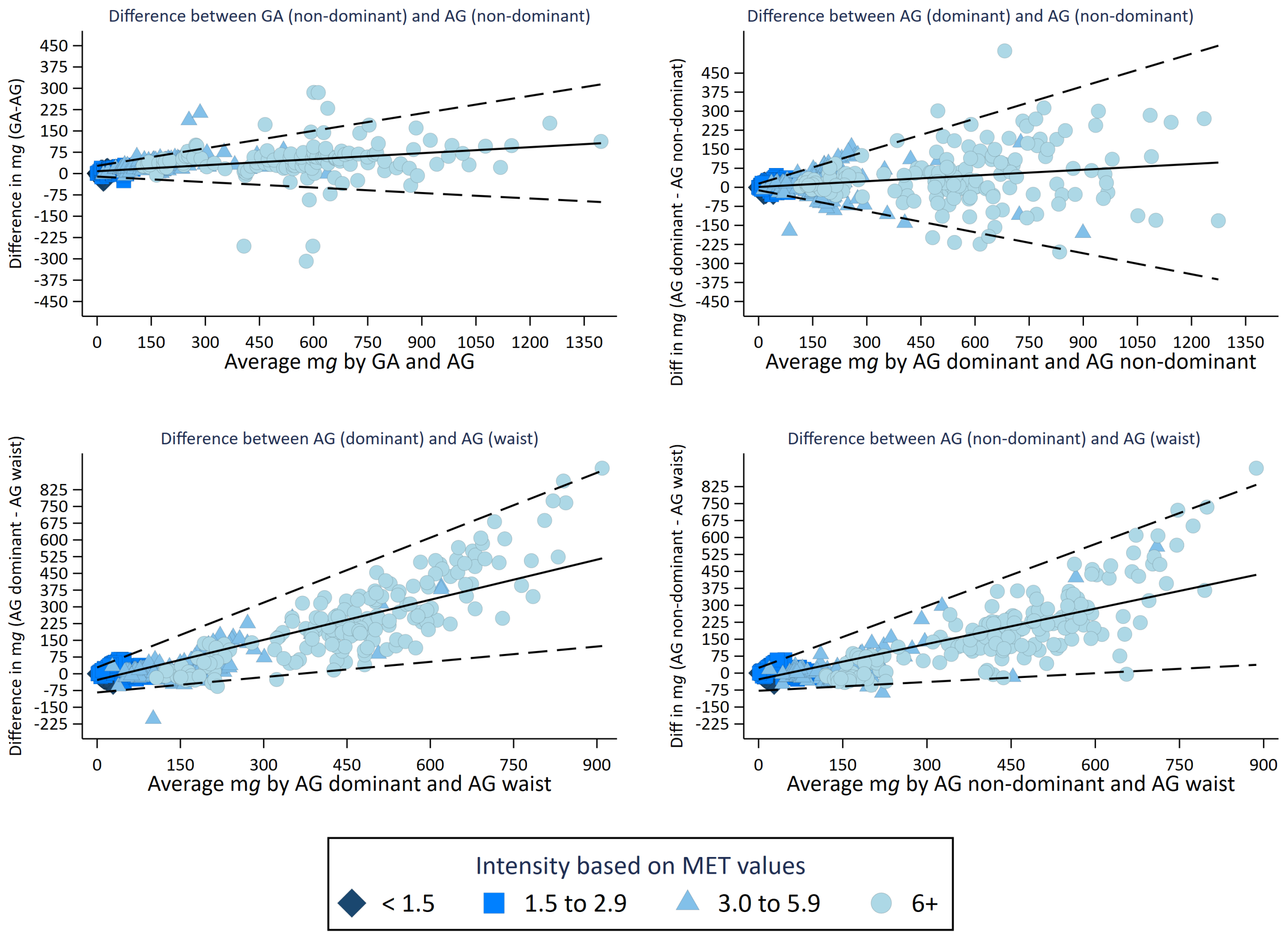

FIGURE 2 Bland–Altman plots of the agreement between acceleration outputs (mg) according to MET values. Solid lines represent the mean bias; dashed lines represent 95% limits of agreement. Mean bias and 95% limits of agreement were estimated using linear regression models based on methods proposed by Bland and Altman. Predictive equations for the mean bias and 95% limits of agreement are described in Table S2. Top left plot represents brand comparison; other plots represent wear-site comparison for AG.

difference between GA and AG non-dominant increased 0.07 units for each increment in mean acceleration; the correlation between difference and mean acceleration was 0.42 (Table S3). The mean difference between dominant and non-dominant AG was 15.6 mg (higher in dominant than in non-dominant; concordance correlation coefficient: 0.97; correlation between difference and mean acceleration: 0.30). The concordance between AG waist and AG in the wrist was 0.80 (mean difference: 51 mg higher) for the non-dominant and 0.75 for dominant (mean difference: 65 mg higher). As observed in Figure 2 and the (Tables S2 and S3; Figure S1), bias, limits of agreement and concordance correlation coefficients between accelerometry outputs varied by activities and intensities. The concordance and absolute mean difference between devices tended to be higher in activities with higher average acceleration, whereas the highest relative mean differences between devices was observed in activities with low average acceleration.

3.2 | Thresholds for inactive and sedentary time

Results from the ROC analyses used to estimate thresholds for inactive time are presented in Table 2 and Figure 3 and for sedentary time are presented in Figure S2 and Table S4. Overall, the area under the curve was similar for wrist monitors regardless of the brand, whereas it was the lowest for the outputs from AG place on the waist. Based on the ROC curves, the acceleration thresholds able to differentiate inactive from active time with the highest sum of sensitivity and specificity were 36 mg for GA non-dominant; 25 mg for AG non-dominant, 30 mg for AG dominant, and 40 mg for AG placed on the waist (Table 2). Overall, the area under the curves were slightly smaller when outputs from sedentary behaviors (sitting and lying) were compared with those from standing, slow walking, brisk walking and running (Figure S2; Table S4).

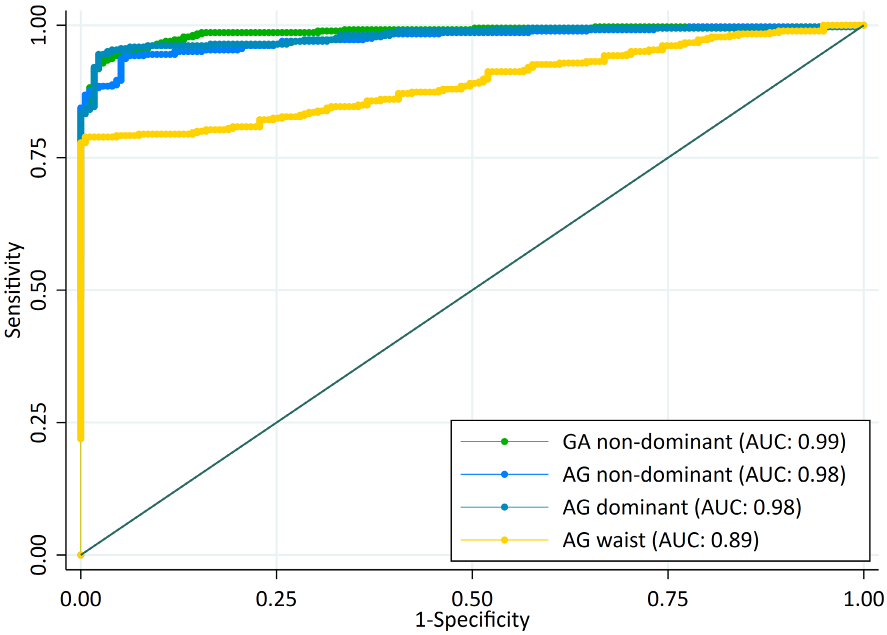

FIGURE 3 Receiver operating characteristics area for definition of inactive time. Inactive time (<1.5 METs) was compared with active time (>1.5 METs). AG, ActiGraph; AUC, Area Under the Curve; GA, GENEActive.

3.3 | Thresholds for moderate and vigorous intensity physical activity

Regression analyses were conducted to predict MET values and intensity of activities based on average acceleration are presented in Figure 4 and Table S5. Linear models from both AG and GA devices at wrist and waist explained a significant proportion of the variance, with adjusted R^2 ranging from 87.6% for AG dominant to 89.8% for GA non-dominant. These were slightly higher than the R^2 observed in the sensitivity analyses that did not include intermittent running or a quadratic polynomial term (Table S5).

3.4 | Proposed intensity thresholds and their accuracy

The proposed absolute intensity thresholds for light (≥1.5 METs), moderate (≥3.0 METs), and vigorous (≥6.0 METs) activities for each device and placement are presented in Table 2. Thresholds for differentiating inactive time from at least light intensity activities ranged from 25 mg for AG placed on the non-dominant wrist to 40 mg for AG placed on the waist. For moderate intensity physical activity, the threshold ranged from 65 mg (AG waist) to 92 mg (GA non-dominant), whereas for vigorous intensity activity optimal thresholds ranged from 190 mg (AG waist) to 283 mg (GA non-dominant). Overall, all thresholds had high sensitivity and specificity in correctly classifying intensities. As demonstrated in Table 3, from all epochs that were classified into a certain intensity, for the wrist at least 73.0% (similar between wrists and between brands) were correctly classified for each intensity, device and placement. For the waist, at least 60.5% were correctly classified for each intensity, device and placement. Overall, from all activities the proposed thresholds for accelerometers placed on the wrist classified as inactive, on average 85% were inactive. The ability of thresholds to accurately classify sedentary behaviors based on posture was approximately 10–20 percentage points lower (Table S6) than the ability to classify sedentary time solely based on energy expenditure (Table 3).

4 | DISCUSSION

This study compared raw triaxial acceleration outputs from two widely used accelerometer brands placed at the wrist and waist in adults performing a range of common daily tasks. Overall, our findings showed that wrist thresholds appear to perform slightly better than waist thresholds for classification of intensities, especially for inactive time. Moreover, this study developed thresholds for raw acceleration that can be utilized in adults aged 20 to 65 years old to reasonably classify physical activities into four categories of absolute intensity. Irrespective of wear-site or device, the inclusion of a quadratic term led to substantial reduction in the vigorous intensity thresholds to the values commonly used in studies that defined intensities of physical activity based on raw acceleration.[3,9]

Overall, high agreement between acceleration outputs from GA and AG brands worn at the same wear-site (non-dominant wrist) was observed. However, agreement varied according to activities and intensities, tending to be better at higher intensities. These findings are

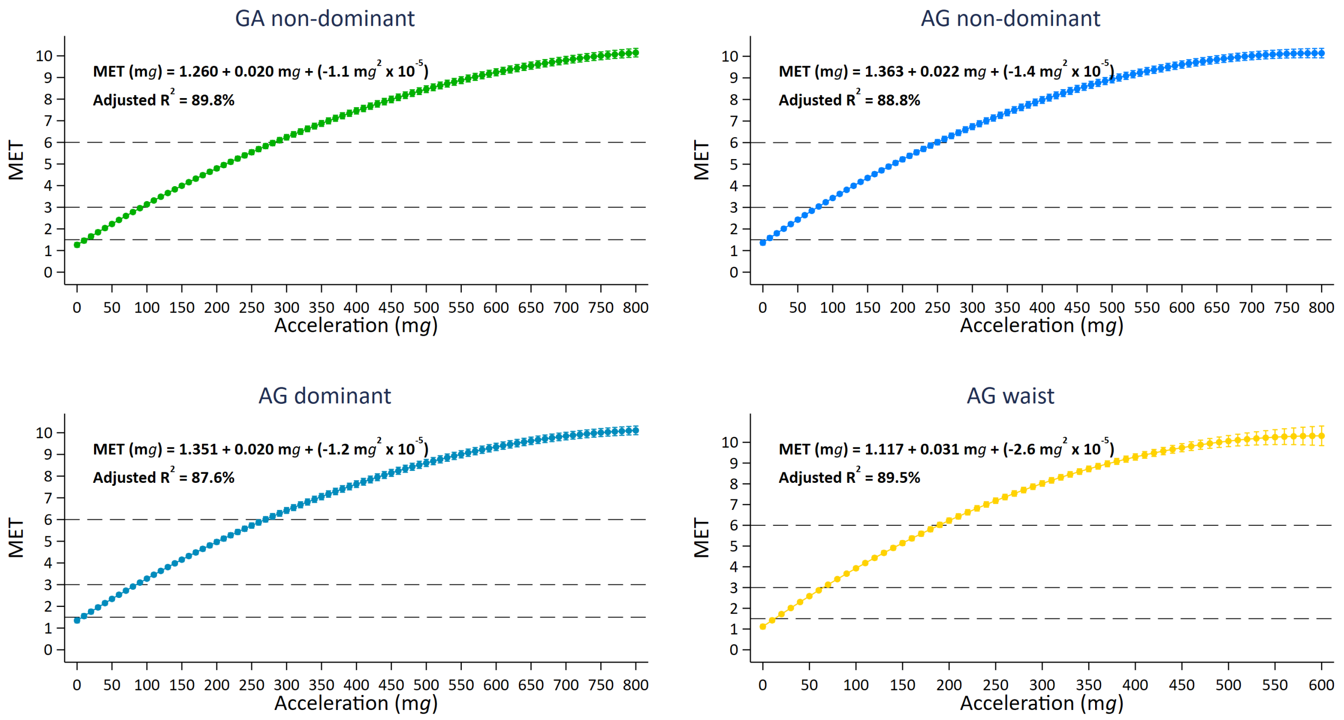

FIGURE 4 Predicted MET values based on raw acceleration outputs from GENEActiv (GA) and ActiGraph (AG). Predicted MET values are based on equation 3 (Table 1).

TABLE 2 Proposed thresholds for absolute intensity based on raw acceleration outputs.

	Threshold in mg	Sensitivity	Specificity	AUC
GA non-dominant				
Light (≥1.5 MET)[a]	36	93	98	0.95
Moderate (≥3MET)[b]	92	93	98	0.98
Vigorous (≥6MET)[b]	283	84	96	0.96
AG non-dominant				
Light (≥1.5 MET)[a]	25	93	95	0.94
Moderate (≥3MET)[b]	78	92	97	0.97
Vigorous (≥6 MET)[b]	249	79	97	0.96
AG dominant				
Light (≥1.5 MET)[a]	30	93	97	0.95
Moderate (≥3MET)[b]	85	91	97	0.97
Vigorous (≥6MET)[b]	270	80	97	0.96
AG waist				
Light (≥1.5 MET)[a]	40	78	100	0.89
Moderate (≥3MET)[b]	65	96	94	0.98
Vigorous (≥6MET)[b]	190	82	92	0.95

[a]Receiver operating characteristics analyses – Youden Index to define empirical optimal threshold.
[b]Regression analyses.

consistent with those of Hildebrand et al.[9] and suggest that accelerometer outputs between brands are comparable for some but not all activities. In our study, we found that GA exhibited higher wrist acceleration than AG, contrary to Hildebrand et al.'s (2014)[9] findings but consistent with Rowlands et al.'s (2018) observations.[7] Furthermore, in a study that compared measures of raw acceleration from wrist-worn GA and AG is 34 adults, Rowlands et al. observed that brands compared well for acceleration outputs higher than 80 mg, but not in lower intensity activities.[6] These findings suggest that caution is needed when comparing estimates of inactivity and sedentary time using a similar absolute intensity threshold between accelerometer brands. It is important to note however, that these differences might be diluted in large population-based studies, in which precision in

TABLE 3 Average confusion matrix indicating the ability of thresholds to accurately classify intensities (% accurately classified).[a] Columns indicate actual intensity, while rows indicate predicted intensity.

	Inactive[b] n (%)	Light n (%)	Moderate n (%)	Vigorous n (%)	Kappa (SE)
GA non-dominant					
Inactive	201 (87.8)	26 (11.4)	2 (0.8)	0 (0.0)	0.78 (0.02)
Light	6 (5.0)	94 (78.3)	20 (16.7)	0 (0.0)	
Moderate	0 (0.0)	6 (4.1)	117 (80.7)	22 (15.2)	
Vigorous	0 (0.0)	0 (0.0)	18 (13.2)	118 (86.7)	
AG non-dominant					
Inactive	206 (85.1)	35 (14.5)	1 (0.4)	0 (0.0)	0.75 (0.02)
Light	12 (8.8)	100 (73.0)	25 (18.3)	0 (0.0)	
Moderate	0 (0.0)	10 (5.8)	130 (75.1)	33 (19.1)	
Vigorous	0 (0.0)	0 (0.0)	14 (10.1)	125 (89.9)	
AG dominant					
Inactive	234 (86.7)	34 (12.6)	2 (0.7)	0 (0.0)	0.76 (0.02)
Light	7 (4.8)	111 (75.5)	29 (19.7)	0 (0.0)	
Moderate	0 (0.0)	11 (6.0)	138 (75.0)	35 (19.0)	
Vigorous	0 (0.0)	0 (0.0)	17 (11.0)	138 (89.0)	
AG waist					
Inactive	245 (68.6)	111 (31.1)	1 (0.3)	0 (0.0)	0.62 (0.02)
Light	1 (2.6)	23 (60.5)	14 (36.8)	0 (0.0)	
Moderate	0 (0.0)	26 (12.8)	146 (71.6)	32 (15.7)	
Vigorous	0 (0.0)	0 (0.0)	27 (15.7)	145 (84.3)	

[a]Interpretation: e.g. AG non-dominant wrist: from all epochs the threshold classified as 'inactive', 85% were inactive (<1.5 METs).

[b]Inactive time was defined based on energy expenditure (<1.5 MET).

population estimate of physical activity is prioritized instead of accurate estimates of individual time spend in specific activities.

The unexpectedly high mean acceleration of AG waist observed when participants were in the lying position was surprising. The mean acceleration was higher than the acceleration reported for the same position in Hildebrand et al. (2014) and the mean acceleration of AG wrist measured in our study. However, the measure of variability of acceleration was quite similar in both studies (standard deviation ~6–7). Furthermore, in Hildebrand et al, the distribution of acceleration during lying activities was largely skewed (coefficient of variation = 281%), whereas in our study, the coefficient of variation was 50%. This could be one of the reasons why the cut-point of sedentary time/inactive time for the waist has a lower sensitivity than the wrist. Although we were unable to identify any significant methodological differences that could account for these disparities, the observed differences in the findings between Hildebrand et al. (2014) and our study emphasize the importance of our and future studies in this area of research.

Findings from our study were similar to those observed in previous studies[9,23,24] that showed differences between AG waist and AG wrist tent to be higher in activities with higher average acceleration, especially for circuit and running activities. To date, the majority of calibration studies to derive thresholds for waist-worn devices were based on brand-specific count-based outcomes,[4] only a few derived intensity thresholds for raw acceleration data.[9] Future research could explore equations for correction estimates and improve comparability between studies and brands.[8]

While some studies have suggested that the thigh is the ideal placement for accelerometers to differentiate sedentary behavior from other activities,[25] it does not capture upper body activities. Further, in general, the wrist and waist are more commonly used in research as these locations offer reasonable accuracy, are less obtrusive, and may be better suited for large-scale population studies where compliance and feasibility are important.[25] Due to the inability to differentiate between postures using magnitude of acceleration alone, it has been suggested that classifying accelerometry thresholds should be used to describe a spectrum of inactive to active data rather than referring to sedentary time, which infers posture.[26]

Deriving thresholds for inactive and sedentary time may require different approaches due to the conceptual difference between these constructs. Given that wrist and waist accelerometers are ideally placed to detect body movement instead of posture,[27] in our study we estimated thresholds for inactivity (based on MET values) and sedentary time (based on posture and MET values). Overall, the accuracy of thresholds was slightly higher for inactive time than for sedentary time, regardless of device brand and wear location. The previous study by Hildebrand and colleagues[11] in adults showed that acceleration of approximately 45 mg (wrist) had 98% of sensitivity and 74% of specificity to distinguish sedentary time (lying and sitting) from non-sedentary time (standing, circuit, slow walking and stepping). When we replicated similar methods used by Hildebrand et al, findings indicated that the optimal wrist thresholds for sedentary time ranged between 12 and 45 mg and presented higher specificity than previously observed.

Absolute intensity thresholds for moderate and vigorous activities identified in our study are lower than those that have been widely used in the literature (100 mg for moderate and 400 mg for vigorous).[9] This might be partially explained by the assumption of non-linearity in the relationship between output acceleration and oxygen consumption, which was demonstrated by the increase in the adjusted-R^2 when a quadratic-term was included in the predictive regression models. Overall, sensitivity, specificity, and area under the curve for absolute intensity thresholds observed in our study were higher than previous indicated in the literature. Although chance cannot be fully disregarded, it is likely that the inclusion of intermittent running in the study protocol, and the use of quadratic polynomial term for acceleration in the regression analyses partially explain this finding. To add robustness to the interpretation of our findings, we also explored different approaches to statistical analyses to identify thresholds, including regression models like those by Hildebrand et al. In these analyses, the thresholds identified for absolute intensity were close to those identified by Hildebrand (Table S7).

More sophisticated approaches can be applied to accelerometer data, e.g. machine learning to categorize activity type.[28] However, the majority of studies that deploy accelerometers still use cut-point approaches to analyze their data.[2,3,18] By calibrating accelerometer output relative to energy expenditure and/or posture, cut-points give biological meaning to accelerometer data.[29] They are easy to apply, provide meaningful outcomes, and enable interpretation of accelerometer data relative to physical activity guidelines. Because of this they have facilitated considerable progress in physical activity and health research. For example, a recent harmonized meta-analysis used cut-point analyses to demonstrate that 30–40 min/day of moderate to vigorous physical activity can attenuate the association between sedentary time and risk of mortality.[30]

However, cut-points have typically been determined in relatively small, laboratory-based samples. As the magnitude of the cut-points developed is heavily influenced by the sample and the activities performed this has led to a proliferation of cut-points in the literature. The time recorded in, e.g., moderate to vigorous physical activity, will differ depending on which cut-points are applied. Migueles et al.[31] recently reported that 8%–96% of their sample of children met physical activity guidelines depending on which cut-point they applied. To advance calibration of accelerometer output, rather than simply adding to the range of cut-points available, Bassett et al.[29] recommended that calibration studies should build on previous evidence and proposed approaches should be evaluated in relation to those already in common use. Thus, a strength of this study was accounting for the non-linear relationship observed between acceleration outputs and MET values and compared the efficacy of the cut-points generated to the widely used cut-points by Hildebrand and colleagues.[9,11]

Limitations of this study should be considered. Although the sample selection was designed to include heterogeneous participants in terms of gender, age and physical activity levels, the convenience sample might limit the generalization of our findings. However, our targeted recruitment ensured that participants with various levels of physical activity were included in order to reflect participants with different fitness levels. The exclusion of participants with chronic conditions is likely to have impacted the thresholds and limit the external validity of the study. The protocol of activities used in this study was composed of nine activities performed in a laboratory environment. Although these activities are likely to be a broad representation of common daily activities, daily physical activity is accrued in more complex patterns with a mixture of light and moderate intensity movement, with episodes of vigorous activity, and are unlikely to be performed following a symmetric, progressive pattern as they were planned for this study. The use of standard METs instead of individualized METs may be considered a limitation of the study. However, this methodological approach was used to allow comparability with the previous study by Hildebrand and colleagues.[9] Moreover, the measured VO2 during lying did not differ substantially from the standard 3.5 mL/kg/min. Lastly, as the derived thresholds may be affected by the activities included in the calibration study, the inclusion of sedentary activities when deriving thresholds for moderate to vigorous activities may have influenced these thresholds. However, this approach was also used by Hildebrand and colleagues.[9] Sensitivity analyses excluding those activities were conducted and did not substantially change the proposed thresholds (data not shown). Our study was designed with several methodological differences compared to the earlier study by Hildebrand, which may introduce additional sources of variability. These were the location, the addition of intermittent running and the quadratic analysis. To allow comparison with the Hildebrand study we also analyzed the data excluding the intermittent running and using the same regression model as Hildebrand. This allowed assessment of the replicability of the Hildebrand study as the only difference for that analysis was the laboratory, participants, and geographical location. Additionally, the potential effect of geographical location on acceleration could impact the generalizability of our findings, making it a potential source of variability. As such, future studies that use consistent measurement protocols and study designs are needed to replicate our findings.

Strengths of this study include the use of a variety of common daily activities which mirrored free-living activities, including intermittent running. Oxygen uptake was assessed using a gold standard method. To predict MET values based on raw triaxial acceleration this study included polynomic quadratic terms, which increased the explanation of the models in 20%. Furthermore, this was the first study of this kind to be conducted in Latin American participants. This can also be considered a strength under the assumption that physiological responses to physical activity may vary across populations with different socioeconomic and ethnical backgrounds, and that error in acceleration estimates may vary by continents.[10]

5 | PERSPECTIVE

Raw accelerometer outputs from GA and AG seem to have high agreement but with limited comparability for the estimation of time spend in low intensity activities. This study reinforces the use of a unique threshold for classifying different intensities of physical activity may be limited when comparing accelerometery outputs from two widely used brands. Moreover, this study provides thresholds for raw accelerometer data that can be used to classify physical activities into absolute intensity categories.

AUTHOR CONTRIBUTIONS

GM conducted the analyses, contributed to data interpretation and wrote the first draft of the manuscript. MM, FR and ICM conceived the study, conducted the data collection and contributed to data interpretation. AR and UE contributed to data interpretation, writing, reviewing and editing. All authors provided critical review and approved the final version of this manuscript.

ACKNOWLEDGEMENTS

We thank all participants for their cooperation during data collection. Open access publishing facilitated by The University of Queensland, as part of the Wiley - The University of Queensland agreement via the Council of Australian University Librarians.

FUNDING INFORMATION

Gregore Mielke is supported by a National Health and Medical Research Council (NHMRC) Investigator Grant (APP2008702; GIM). Alex Rowlands is supported by the NIHR Leicester Biomedical Research Centre, and the NIHR Applied Research Collaboration (ARC) East Midlands. Inacio Crochemore-Silva is supported by the Brazilian National Research Council (CNPq). The views expressed are those of the authors and not necessarily those of the NHMRC, NHS, NIHR, Department of Health, or CNPq.

CONFLICT OF INTEREST STATEMENT

All authors have no conflict of interest to declare. The conclusions of the present study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Gregore Iven Mielke https://orcid.org/0000-0002-3043-2715
Márcio de Almeida Mendes https://orcid.org/0000-0002-5590-5750
Ulf Ekelund https://orcid.org/0000-0003-2115-9267
Alex V. Rowlands https://orcid.org/0000-0002-1463-697X
Felipe Fossati Reichert https://orcid.org/0000-0002-0951-9875
Inacio Crochemore-Silva https://orcid.org/0000-0001-5390-8360

REFERENCES

1. Strath SJ, Kaminsky LA, Ainsworth BE, et al. Guide to the assessment of physical activity: clinical and research applications: a scientific statement from the American Heart Association. *Circulation*. 2013;128(20):2259-2279. doi:10.1161/01.cir.0000435708.67487

2. Migueles JH, Cadenas-Sanchez C, Ekelund U, et al. Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. *Sports Med*. 2017;47(9):1821-1845. doi:10.1007/s40279-017-0716-0
3. Brady R, Brown WJ, Hillsdon M, et al. Patterns of accelerometer-measured physical activity and health outcomes in adults: a systematic review. *Med Sci Sports Exerc*. 2022;54:1155-1166. doi:10.1249/MSS.0000000000002900
4. Migueles JH, Rowlands AV, Huber F, et al. GGIR: a research community–driven open source R package for generating physical activity and sleep outcomes from multi-day raw accelerometer data. *Journal for the Measurement of Physical Behaviour*. 2019;2(3):188-196. doi:10.1123/jmpb.2018-0063
5. Wijndaele K, Westgate K, Stephens SK, et al. Utilization and harmonization of adult accelerometry data: review and expert consensus. *Med Sci Sports Exerc*. 2015;47(10):2129-2139. doi:10.1249/MSS.0000000000000661
6. Rowlands AV, Yates T, Davies M, et al. Raw accelerometer data analysis with GGIR R-package: does accelerometer brand matter? *Med Sci Sports Exerc*. 2016;48(10):1935-1941. doi:10.1249/MSS.0000000000000978
7. Rowlands AV, Mirkes EM, Yates T, et al. Accelerometer-assessed physical activity in epidemiology: are monitors equivalent? *Med Sci Sports Exerc*. 2018;50(2):257-265. doi:10.1249/MSS.0000000000001435
8. Rowlands AV, Fraysse F, Catt M, et al. Comparability of measured acceleration from accelerometry-based activity monitors. *Med Sci Sports Exerc*. 2015;47(1):201-210. doi:10.1249/MSS.0000000000000394
9. Hildebrand M, VT VANH, Hansen BH, Ekelund U. Age group comparability of raw accelerometer output from wrist- and hip-worn monitors. *Med Sci Sports Exerc*. 2014;46(9):1816-1824. doi:10.1249/MSS.0000000000000289 [published Online First: 2014/06/03].
10. van Hees VT, Fang Z, Langford J, et al. Autocalibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: an evaluation on four continents. *J Appl Physiol (1985)*. 2014;117(7):738-744. doi:10.1152/japplphysiol.00421.2014 [published Online First: 2014/08/12].
11. Hildebrand M, Hansen BH, van Hees VT, Ekelund U. Evaluation of raw acceleration sedentary thresholds in children and adults. *Scand J Med Sci Sports*. 2017;27(12):1814-1823. doi:10.1111/sms.12795
12. Doherty A, Jackson D, Hammerla N, et al. Large scale population assessment of physical activity using wrist worn accelerometers: the UK biobank study. *PLoS One*. 2017;12(2):e0169649. doi:10.1371/journal.pone.0169649
13. Belcher BR, Wolff-Hughes DL, Dooley EE, et al. US population-referenced percentiles for wrist-worn accelerometer-derived activity. *Med Sci Sports Exerc*. 2021;53(11):2455-2464. doi:10.1249/MSS.0000000000002726
14. da Silva IC, van Hees VT, Ramires VV, et al. Physical activity levels in three Brazilian birth cohorts as assessed with raw triaxial wrist accelerometry. *Int J Epidemiol*. 2014;43(6):1959-1968. doi:10.1093/ije/dyu203
15. Ding D, Mielke GI, Silva ICM, et al. Prenatal and birth predictors of objectively measured physical activity and sedentary time in three population-based birth cohorts in Brazil. *Sci Rep*. 2020;10(1):786. doi:10.1038/s41598-019-57070-x
16. Mielke GI, Menezes AMB, Mielke GI, et al. Associations between device-measured physical activity and cardiometabolic health in the transition to early adulthood. *Med Sci Sports Exerc*. 2021;53(10):2076-2085. doi:10.1249/MSS.0000000000002696 [published Online First: 2021/05/10].
17. Mendes MA, da Silva I, Ramires V, et al. Metabolic equivalent of task (METs) thresholds as an indicator of physical activity intensity. *PLoS One*. 2018;13(7):e0200701. doi:10.1371/journal.pone.0200701
18. Bull FC, Al-Ansari SS, Biddle S, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *Br J Sports Med*. 2020;54(24):1451-1462. doi:10.1136/bjsports-2020-102955
19. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res*. 1999;8(2):135-160. doi:10.1177/096228029900800204
20. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307-310. [published Online First: 1986/02/08].
21. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989;45(1):255-268. [published Online First: 1989/03/01].
22. Tremblay MS, Aubert S, Barnes JD, et al. Sedentary behavior research network (SBRN) – terminology consensus project process and outcome. *Int J Behav Nutr Phys Act*. 2017;14(1):75. doi:10.1186/s12966-017-0525-8
23. Rosenberger ME, Haskell WL, Albinali F, et al. Estimating activity and sedentary behavior from an accelerometer on the hip or wrist. *Med Sci Sports Exerc*. 2013;45(5):964-975. doi:10.1249/MSS.0b013e31827f0d9c
24. Migueles JH, Cadenas-Sanchez C, Rowlands AV, et al. Comparability of accelerometer signal aggregation metrics across placements and dominant wrist cut points for the assessment of physical activity in adults. *Sci Rep*. 2019;9(1):18235. doi:10.1038/s41598-019-54267-y
25. Bakrania K, Yates T, Rowlands AV, et al. Intensity thresholds on raw acceleration data: euclidean norm minus one (ENMO) and mean amplitude deviation (MAD) approaches. *PLoS One*. 2016;11(10):e0164045. doi:10.1371/journal.pone.0164045
26. Sedentary Behaviour Research Network. Letter to the editor: standardized use of the terms "sedentary" and "sedentary behaviours". *Appl Physiol Nutr Metab*. 2012;37(3):540-542. doi:10.1139/h2012-024 Epub 2012 Apr 27. PMID: 22540258.
27. Chen KY, Bassett DR Jr. The technology of accelerometry-based activity monitors: current and future. *Med Sci Sports Exerc*. 2005;37(11 Suppl):S490-S500. doi:10.1249/01.mss.0000185571.49104.82
28. Greenwood-Hickman MA, Nakandala S, Jankowska MM, et al. The CNN hip accelerometer posture (CHAP) method for classifying sitting patterns from hip accelerometers: a validation study. *Med Sci Sports Exerc*. 2021;53(11):2445-2454. doi:10.1249/MSS.0000000000002705
29. Bassett DR Jr, Rowlands A, Trost SG. Calibration and validation of wearable monitors. *Med Sci Sports Exerc*. 2012;44(1 Suppl 1):S32-S38. doi:10.1249/MSS.0b013e3182399cf7
30. Ekelund U, Tarp J, Fagerland MW, et al. Joint associations of accelero-meter measured physical activity and sedentary time with all-cause mortality: a harmonised meta-analysis in more than 44 000 middle-aged and older individuals. *Br J Sports Med*. 2020;54(24):1499-1506. doi:10.1136/bjsports-2020-103270

31. Migueles JH, Cadenas-Sanchez C, Tudor-Locke C, et al. Comparability of published cut-points for the assessment of physical activity: implications for data harmonization. *Scand J Med Sci Sports*. 2019;29(4):566-574. doi:10.1111/sms.13356

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Mielke GI, de Almeida Mendes M, Ekelund U, Rowlands AV, Reichert FF, Crochemore-Silva I. Absolute intensity thresholds for tri-axial wrist and waist accelerometer-measured movement behaviors in adults. *Scand J Med Sci Sports*. 2023;33:1752-1764. doi:10.1111/sms.14416