International Journal of EXERCISE SCIENCE

Original Research

Measurement of Physical Activity with Wrist-Worn ActiGraph GT3X+ in Older Women

MICHAL T. SMITH†[1], ERIN E. KISHMAN†[1], R. GLENN WEAVER‡[1], JENNIFER R. O'NEILL‡[1], and XUEWEN WANG‡[1]

[1]Department of Exercise Science, Arnold School of Public Health, University of South Carolina, Columbia, South Carolina, USA

*Denotes undergraduate student author, †Denotes graduate student author, ‡Denotes professional author

ABSTRACT

International Journal of Exercise Science 15(7): 1538-1553, 2022. Higher wear compliance has been seen with wrist placed accelerometers versus hip placed. Performance of wrist placed ActiGraph GT3X+ accelerometer (GT3X+, ActiGraph LLC, Pensacola, FL) in assessing physical activity (PA) remains unclear. PURPOSE: This study examined GT3X+'s performance in measuring PA energy expenditure (PAEE) and classifying PA intensity in older women. METHODS: Women [n = 89, age = 65.6 (4.3)] wore GT3X+ and SenseWear Armband Mini (SWAM, BodyMedia Inc. Pittsburgh, PA) for 2 weeks. Concurrently, doubly labeled water (DLW) determined total daily energy expenditure (TDEE). Resting energy expenditure (REE) was determined by Indirect Calorimetry. Data was processed using manufacturer-provided software. Bivariate correlations, Intra Class Correlations, and Bland-Altman plots were performed to evaluate agreement between GT3X+ and criterion measures for sedentary time, light and moderate-to-vigorous PA (determined by SWAM) and PAEE (determined by SWAM and by DLW and REE). Epoch-by-epoch analysis evaluated discrepancy and agreement of PA intensity classification between GT3X+ and SWAM. RESULTS: For PAEE, GT3X+ showed moderate correlations with criterion measures (r = 0.413, 0.400 with SWAM; r = 0.564, 0.501 with DLW and REE), but Bland-Altman plots showed large variability. When estimating time spent in PA intensity, GT3X+ underestimated sedentary time and overestimated PA intensity compared to SWAM. During epoch-by-epoch analysis, GT3X+ misclassified light intensity PA as moderate-to-vigorous PA 72% of the time. Counts per minute showed strong correlations with criterion measures (r = 0.68, 0.625 for SWAM and DLW and REE respectively). CONCLUSION: Current equations and cut points do not provide accurate measures of PA with wrist-worn GT3X+ in older women.

KEY WORDS: Accelerometer, energy expenditure, postmenopausal women, wrist accelerometry, exercise.

INTRODUCTION

Being able to accurately measure physical activity intensity and energy expenditure (EE) is crucial to understanding the role physical activity plays in lowering the risk for disease (8, 15, 16, 26, 30). The combination of indirect calorimetry and the doubly labeled water method (DLW) is the gold standard method of measuring physical activity energy expenditure (PAEE); however, these measures have limitations such as cost and cumbersome equipment, making it difficult to measure day-to-day PAEE in free-living conditions (25).

Accelerometry is a widely used method of measuring sedentary time, physical activity, and EE because it poses many benefits over the aforementioned methods. Accelerometers are easy to use, low in cost when compared to the DLW and are more accurate than self-report measures (22). Furthermore, unlike indirect calorimetry, they allow measurement of free-living PAEE (11). However, the validity of the available research grade accelerometers varies across brands, models, and placements (5). The SenseWear Armband Mini (SWAM, BodyMedia Inc. Pittsburgh, PA, USA) monitor is a previously validated multisensor monitor that provides estimates of PAEE, sedentary time, and active time. The SWAM is no longer being produced, therefore increasing the need for other valid and unobtrusive methods of estimating EE.

The GT3X+ ActiGraph accelerometer (GT3X+, ActiGraph LLC, Pensacola, FL) is a small triaxial accelerometer that can be worn on the hip or the wrist. Several studies (1, 4, 10) have shown acceptable validity of the GT3X+ when placed on the hip, compared to indirect calorimetry. This has been seen throughout a range of activity modalities and intensities (± 10% equivalence zone) (1). However, the placement presents challenges with under detection of activities with substantial arm movement (4). Furthermore, the wrist placement may be the more comfortable placement for sleeping and daytime wear. Previous research has suggested that the wrist placement is effective in improving participant compliance, which would make the wrist worn GT3X+ an ideal choice for many researchers (2, 9, 19, 29, 31).

Few studies have aimed to evaluate the validity of the wrist-worn GT3X+. McMinn et al (17) found that PAEE estimates from the GT3X+ highly correlated with indirect calorimetry under controlled walking conditions (r = 0.72). However, Bland-Altman plots revealed that the GT3X+ significantly underestimated PAEE during slow walking (mean difference of 1.22 kcal·min^{-1}) and significantly overestimated PAEE during fast walking (mean difference of 0.96 kcal·min^{-1}) (17). Hildebrand et al (12) compared the raw accelerometer outputs of the GT3X+ from both the hip and wrist placement in 30 adults during a range of activity intensities. ICCs between different placements of the GT3X+ for raw accelerations was 0.905 with a 95% confidence interval (CI) of 0.903-0.907. Furthermore, the authors mention that taking into account the less obtrusive nature of the wrist placement, the ability to improve wear compliance makes the wrist placement a viable option (12).

The few studies that aimed to evaluate the validity of wrist-worn GT3X+ used controlled intensities (12, 17); there is a lack of studies evaluating the performance of the GT3X+ in free-living conditions where activity patterns are more complex. This study aims to examine the performance of the wrist-worn GT3X+ in measuring PAEE and classifying activity intensity in a healthy non-obese population of older women. Specifically, we aimed to evaluate discrepancy and agreement of the wrist-worn GT3X+ in measuring PAEE when compared to PAEE estimates from the SWAM and PAEE estimates derived from DLW-measured TDEE and indirect calorimetry-determined resting EE (REE) in a population of older women. Second, we aimed to evaluate the discrepancy and agreement of the wrist-worn GT3X+ in measuring time spent in sedentary activity, light intensity activity and MVPA when compared to time in intensity estimates determined by the SWAM. Furthermore, we aimed to use epoch-by-epoch analysis to quantify the discrepancies in activity intensity classification between the GT3X+ and the SWAM.

METHODS

Baseline data collected in the Women's Energy Expenditure in Walking Programs (WeWalk) Study (ClinicalTrials.gov identifier: NCT01722136) were used for this study (3, 32). The WeWalk Study was a 16-week randomized controlled trial to investigate the effects of two different doses of moderate intensity exercise on energy expenditure in inactive older women. Study procedures were reviewed and approved by the University of South Carolina Institutional Review Board in Columbia, South Carolina. All participants signed a written informed consent prior to participation in the study. This research was carried out fully in accordance to the ethical standards of the International Journal of Exercise Science (18).

Participants

Participant inclusion criteria included age (60-75 years), body mass index (18-30 $kg \cdot m^{-2}$), self-reported stable weight (± 3%) for the past three months, physically inactive (less than 20 min, three times per week of structured exercise) for the past three months, nonsmoking for the past year, and able to walk on a treadmill (3, 32). Exclusion criteria included self-reported serious cardiovascular, metabolic, or respiratory diseases, or other conditions that might affect protocol adherence, exercise safety, or be aggravated by exercise. Participants were also excluded if they were taking medications known to affect exercise performance or metabolism or reported excess caffeine use (> 500 $mg \cdot day^{-1}$).

Protocol

The DLW procedure was published more in depth previously (32). Briefly, the procedure included collecting urine samples over a 14-day period after participants consumed an oral dose of premixed $^2H_2{}^{18}O$. The enrichment of ^{18}O and ^2H in the urine samples was analyzed using the isotope ratio mass spectrometry, and TDEE was calculated following the standard procedures established at the Pennington Biomedical Research Center.

Resting energy expenditure was measured on the last day of the DLW period after an overnight fast using indirect calorimetry performed under a ventilated hood. All measurements took place in the morning between 600 and 800h and at least 24 hours after the last bout of any structured exercise. The procedure was described previously (32). To calculate PAEE, the thermic effect of food was assumed to account for 10% of TDEE (21, 28). PAEE was calculated as TDEE*0.9 – REE.

Except for during water activities, participants were instructed to wear the SWAM for the entire 14-day period that TDEE was measured by DLW. The SWAM was worn on the upper left arm, over the participants' triceps. The manufacturer provided software (SenseWear Professional 8.0, BodyMedia, Inc.) was used to calculate EE and activity intensity for the SWAM. The software uses data from the monitor's sensors (heat flux, galvanic skin response, skin temperature, and near body ambient temperature) and individual information (age, sex, height, weight, smoking, and handedness) to give estimates of EE for each minute of wear time. EE estimates for each minute are then converted to METs which are then used to classify activity intensity. METs for sedentary, light, moderate, and vigorous activity are 1-1.5, 1.6-2.9, 3-5.9, and ≥ 6 METs respectively (33).

The GT3X+ was worn on the non-dominant wrist during the same 14-day period TDEE was measured. Participants were instructed to wear the device for the entire 14-day period and to go about their normal weekly routines. The manufacturer provided software (ActiLife 6.9.5, ActiGraph, LLC) was used to calculate PAEE and activity intensity. PAEE equations and activity intensity cut points were chosen based on the demographics of the populations they had been validated in. Two different equations were used to calculate PAEE: *Freedson VM3 2011* (24) (VM3) and *Freedson VM3 Combination 2011* (24) (VM3 combo). A total of 3 sets of cut points were used to classify activity intensity: *Freedson Adult VM3 2011* (24) (Freedson), *Keadle Women's Health VM 2014* (14) (Keadle) and *Santos-Lozano Older Adults 2013 (23)* (Santos). The Freedson and Keadle cut points were provided by the ActiLife Software while the Santos cut points were manually entered into the program. All cut points used were based on the 3-axis vector magnitude (VM). Details of the cut points are listed in Table 1. The participants were inactive; to make comparison of cut points easier, moderate, vigorous, and very vigorous were combined into one category (MVPA). Minute-by-minute raw data were exported to Excel.

Table 1. Cut Points Used for Estimation of Activity Intensity

Cut Point	Short name	Sedentary	Light	Moderate	Vigorous	Very Vigorous
Freedson Adult VM3 2011	Freedson		0-2689	2690-6166	6167-9642	9643 +
Keadle Women's Health VM 2014	Keadle	0-199	200-2689	2690 +		
Santos-Lozano Older Adults 2013	Santos		0-2750	2751-9359	9360 +	

Values represent vector magnitude counts per minute.

Statistical Analysis

Participants with at least five days of wear were included in the analysis. Days with at least 22 hours of wear time for both monitors were included in the analysis. PAEE and time spent on specific PA intensity were tested for normal distribution. Pearson correlation (or Spearman correlation for models with non-normal variables) and linear regression analyses (any non-normally distributed variables were used as independent variables) were performed to assess the associations of PAEE estimates between the GT3X+ and DLW and IC-determined PAEE, and between GT3X+ and SWAM. PAEE determined by GT3X+ included the VM3 and VM3 combo equation. The same analyses were used to assess association of time estimates of activity intensities between the GT3X+, determined by the GT3X+ included Freedson, Keadle, and

Santos-Lozano cut points, and the SWAM. Furthermore, we examined the associations of daily vector magnitude counts per minute (CPM) from the GT3X+ with PAEE estimates from the SWAM and DLW and IC.

Intra Class Correlations (ICCs) were determined to assess reliability using the two-way mixed model, single measures, and absolute agreement except for the CPM and PAEE comparison in which consistency agreement was used. Bland-Altman plots were used to determine mean bias, trends, and the degree of agreement within the 95% confidence intervals of PAEE estimates between GT3X+ and DLW and IC-determined PAEE and between GT3X+ and SWAM, estimates of time in each intensity between the SWAM and the three different sets of GT3X+ cut points, between GT3X+ CPM and DLW and IC-determined PAEE, and GT3X+ CPM and SWAM PAEE. Analysis was performed using SAS OnDemand (SAS Institute, Cary, NC), except for analysis of ICCs which were ran in SPSS (version 20, Armonk, NY).

The entire epoch-by-epoch analysis was performed in Excel. All GT3X+ data were reintegrated into 60-second epochs in order to match the SWAM data. The Keadle cut points were chosen for this analysis because they provide estimates of sedentary time, making them more comparable to the SWAM. They were applied to the minute-by-minute GT3X+ data and each epoch was given a code to identify its activity intensity. Sedentary activity was coded as 0, light activity was coded as 1 and MVPA was coded as 2. The SenseWear software automatically coded each minute of wear by activity intensity based on METs. For each participant, the GT3X+ data were matched by time with the SWAM data. For each epoch, the GT3X+ was evaluated on whether it agreed with the SWAM on activity intensity. The percent of epochs that agreed was calculated overall and for each of the activity intensities. Furthermore, we identified discrepancies between the two monitors. This was done by calculating what percentage of epochs did not agree and what intensity they were incorrectly classified as. This was done for each participant. Means and standard deviations for all participants were calculated for the following: overall agreement, sedentary agreement, light agreement, MVPA agreement, sedentary classified as light, sedentary classified as MVPA, light classified as sedentary, light classified as MVPA, MVPA classified as sedentary, and MVPA classified as light.

RESULTS

Out of the 89 participants who provided baseline data, the mean age was 65.6 years and the mean BMI was 25.6 kg/m^2. The majority of participants were White/European American (84.3%) and completed 4 years or more of college (63.64%). Annual income ranged from $10,000-19,000 to $80,000+ with 44.58% of the women falling in the $80,000+ category. The majority of women were either employed for wages (38.64%) or retired (51.14%). Means and standard deviations for daily PAEE and time spent in activity intensities are shown in Table 2.

Table 2. Descriptive Statistics of Activity Measures

Measures	*n*	Mean ± SD
PAEE (kilocalories/day)		
SWAM	87	1160.1 ± 1088.1
DLW and IC-derived	72	630.7 ± 202.1
GT3X+ (VM3)	87	1414.4 ± 509.8
GT3X+ (VM3 Combo)	87	1552.8 ± 509.3
CPM		
GT3X+	87	1591.1 ± 501.5
Sedentary Time (minutes/day)		
SWAM	85	1116.1 ± 120.5
GT3X+ (Keadle)	85	866.3 ± 110.5
Time in Light Intensity (minutes/day)		
SWAM	85	261.7 ± 103.1
GT3X+ (Freedson)	85	1130.5 ± 85.8
GT3X+ (Keadle)	85	264.2 ± 50.1
GT3X+ (Santos)	85	1138.6 ± 84.8
Time in MVPA (minutes/day)		
SWAM	85	39.1 ± 27.0
GT3X+ (Freedson)	85	308.2 ± 87.1
GT3X+ (Keadle)	85	308.2 ± 87.1
GT3X+ (Santos)	85	300.2 ± 86.1

Physical Activity Energy Expenditure (PAEE), SenseWear Armband Mini monitor (SWAM), Doubly Labeled Water Method (DLW), Indirect Calorimetry (IC), ActiGraph GT3X+ accelerometer (GT3X+), Vector Magnitude Counts per Minute (CPM), Moderate to Vigorous Physical Activity (MVPA).

Associations and ICCs between measures of PAEE are shown in Table 3. According to the linear regression models and correlations, the VM3 and VM3 combo equations showed similar results when compared to both the DLW method and the SWAM. Both correlations of VM3 and VM3 combo with DLW and IC-determined PAEE were greater than 0.50, indicating large effect size, while their correlations with SWAM-determined PAEE were moderate.

Table 3. Associations and ICCs between Measures of PAEE

Independent	Dependent	*n*	Regression	*P*-value	*r*	ICCs
SWAM	VM3	85	0.15 (0.06, 0.25)	0.0024	0.413*	0.240 (0.038, 0.425)
SWAM	VM3 combo	85	0.16 (0.06, 0.26)	0.0015	0.400*	0.235 (0.035, 0.420)
DLW and IC	VM3	70	1.23 (0.72, 1.75)	< 0.001	0.564*	0.112 (-0.072, 0.338)
DLW and IC	VM3 combo	70	1.21 (0.70, 1.72)	< 0.001	0.501	0.088 (-0.060, 0.293)

Data presented as regression coefficient (95% CI) and ICCs (95% CI). *Indicates Spearman correlation.

However, both equations had poor reliability when compared to the SWAM and even poorer when compared to DLW and indirect calorimetry, indicating poor resemblance of data between PAEE determined by GT3X+ (either equation) and the criterion measures (by SWAM and DLW and IC). Bland-Altman plots in Figures 1a and 1b show large limits of agreement and an overestimation of PAEE when compared to DLW and IC, indicating poor agreement. Figures 1c and 1d also shows large limits of agreement for both equations when compared to the SWAM. Furthermore, when compared to the SWAM, the VM3 and VM3 combo equations tend to overestimate PAEE at low levels of activity and underestimate PAEE at high levels of activity.

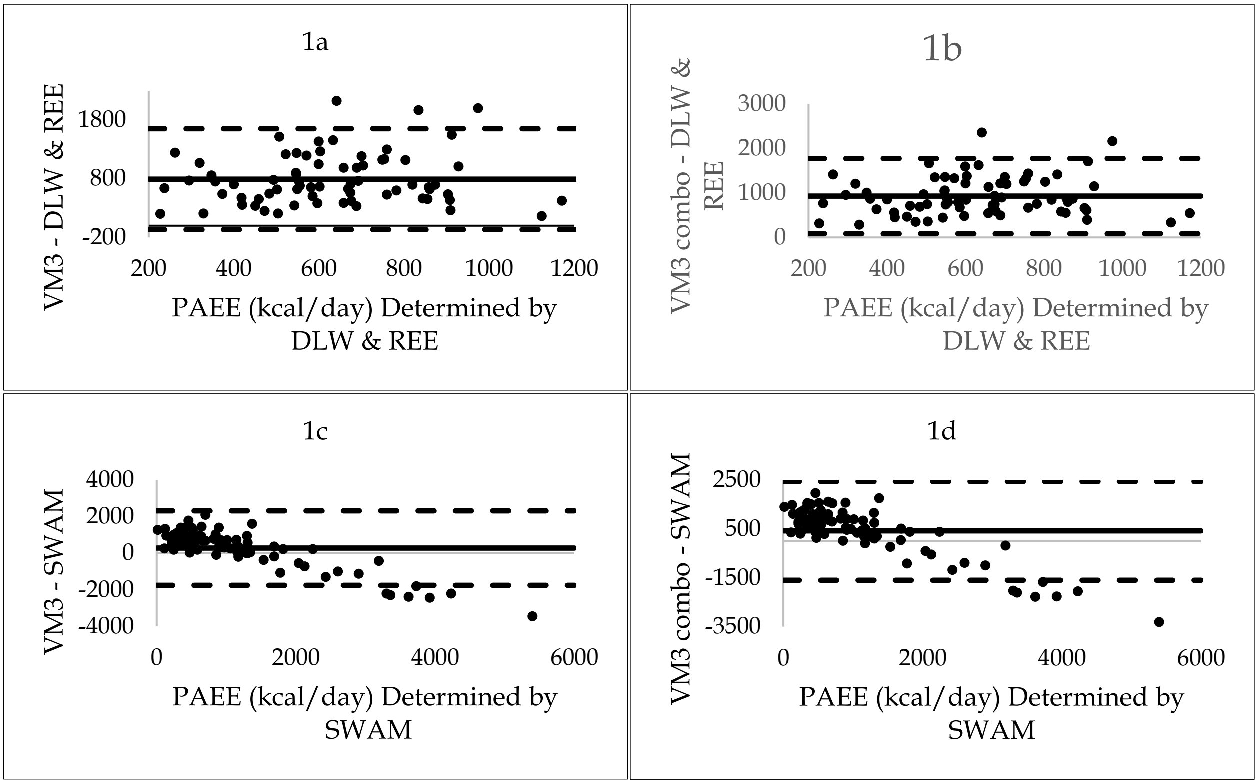

Figure 1. Bland-Altman plots for agreement between Freedson physical activity energy expenditure (PAEE) equations and criterion methods of measuring PAEE. 1a) Difference between Freedson VM3 2011 equation (VM3) and doubly labeled water method (DLW) and Resting energy expenditure (REE). 1b) Difference between Freedson VM3 combination 2011 equation (VM3 combo) and DLW and REE. 1c) Difference between VM3 and SenseWear Armband Mini monitor (SWAM). 1d) Difference between VM3 combo and SWAM.

Associations between SWAM and DLW and indirect calorimetry determined PAEE, and SWAM and DLW TDEE can be found in Table 4. Correlation coefficients were moderate for PAEE and high for TDEE. ICCs were low for both relationships indicating poor reliability. Furthermore, Bland-Altman plots for the agreement between these measures can be found in Figure 2.

Table 4. Associations of PAEE and TDEE determined by Criterion Measures

Independent	Dependent	*n*	Regression	*P*-value	*r*	ICCs
PAEE						
SWAM	DLW	71	0.05(0.009-0.089)	0.0170	0.413*	0.076 (-0.107 – 0.270)
TDEE						
SWAM	DLW	70	0.21(0.16-0.26)	< 0.001	0.698*	0.010 (-0.007-0.047)

Data presented as regression coefficient (95% CI) and ICCs (95% CI). *Indicates Spearman correlation.

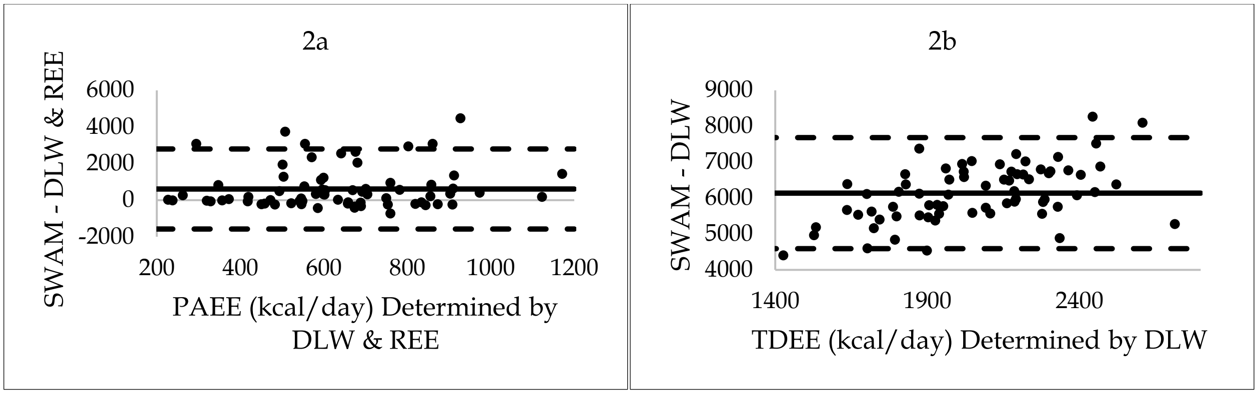

Figure 2. Bland-Altman plots for agreement between criterion measures of physical activity energy expenditure (PAEE) and total daily energy expenditure (TDEE). a) PAEE: Difference between SenseWear Armband Mini monitor (SWAM) and doubly labeled water method (DLW) and Resting energy expenditure (REE). b) TDEE: Difference between SWAM and DLW.

Associations and ICCs between estimates of time spent in activity intensity are seen in Table 5. The Keadle cut points were the only cut points to produce estimates of sedentary time while the other cut points only produced estimates of light intensity activity and MVPA. Pearson's correlation and linear regression coefficients between Keadle sedentary time and SWAM sedentary time were high, while ICCs were low. A Bland-Altman plot for sedentary time is shown in Figure 3. Large limits of agreement are seen along with a trend of consistently underestimating time spent in sedentary time compared to SWAM.

Table 5. Associations and ICCs between Measures of Time in Intensity

Independent	Dependent	*n*	Regression	*P*-value	*r*	ICCs
Sedentary						
SWAM	Keadle	85	0.65 (0.50, 0.79)	< 0.001	0.705	0.211 (-0.053, 0.551)
Light						
SWAM	Freedson	85	-0.58 (-0.71, -0.45)	< 0.001	-0.701	-0.016 (-0.020 0.020)
SWAM	Keadle	85	0.15 (0.05, 0.25)	0.0037	0.329*	0.247 (0.035, 0.437)
SWAM	Santos	85	-0.57 (-0.70, -0.44)	< 0.001	-0.697	-0.016 (-0.019, 0.020)
MVPA						
SWAM	Freedson	85	2.08 (1.55, 2.62)	< 0.001	0.676*	0.038 (-0.026 0.154)
SWAM	Keadle	85	2.08 (1.55, 2.62)	< 0.001	0.676*	0.038 (-0.026, 0.154)
SWAM	Santos	85	2.07 (1.54, 2.60)	< 0.001	0.675*	0.039 (-0.027, 0.160)

Data presented as regression coefficient (95% CI) and ICCs (95% CI). *Indicates Spearman Correlation.

As seen in Table 5, when estimating time spent in light intensity activity, the Keadle cut points performed better on the linear regression and correlation analyses because they were the only cut points to produce positive coefficients. Although the Keadle cut points performed better than the other cut points, the regression coefficients and correlation coefficients they produced would still be considered low. ICCs for light intensity activity can also be seen in Table 5. Although, highest for the Keadle cut points, ICCs indicated poor reliability for all cut points when compared to the SWAM. Bland-Altman plots for light activity are shown in Figure 4a, 4b, and 4c. Again, the best agreement in measuring light intensity activity is seen between the two devices when the Keadle cut points are used.

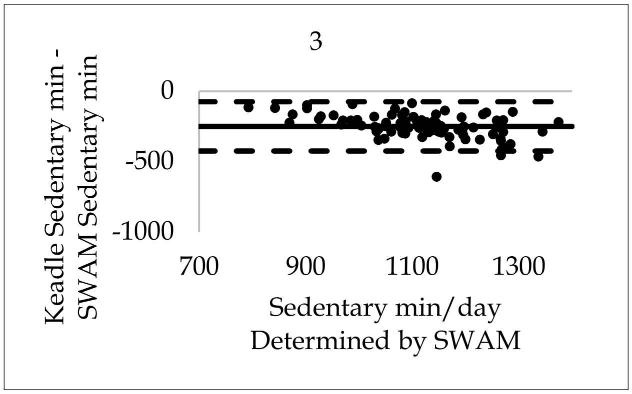

Figure 3. Bland-Altman plot for agreement of time spent in sedentary activity time between Keadle Women's Health VM 2014 cut points (Keadle) and SenseWear Armband Mini monitor (SWAM).

All cut points produced similar estimates of time spent in MVPA. Therefore, no set of cut points outperformed another regarding regression and correlations coefficients for time spent in MVPA when compared to the SWAM (Table 5), and all correlations for all cut points are high. Similarly, based on ICCs, no cut point seems to be largely outperforming the other in reliability of measuring MVPA. Bland-Altman plots for MVPA are shown in Figure 5a, 5b, and 5c. These plots show a trend of overestimating time spent in MVPA that increases as intensity level increases.

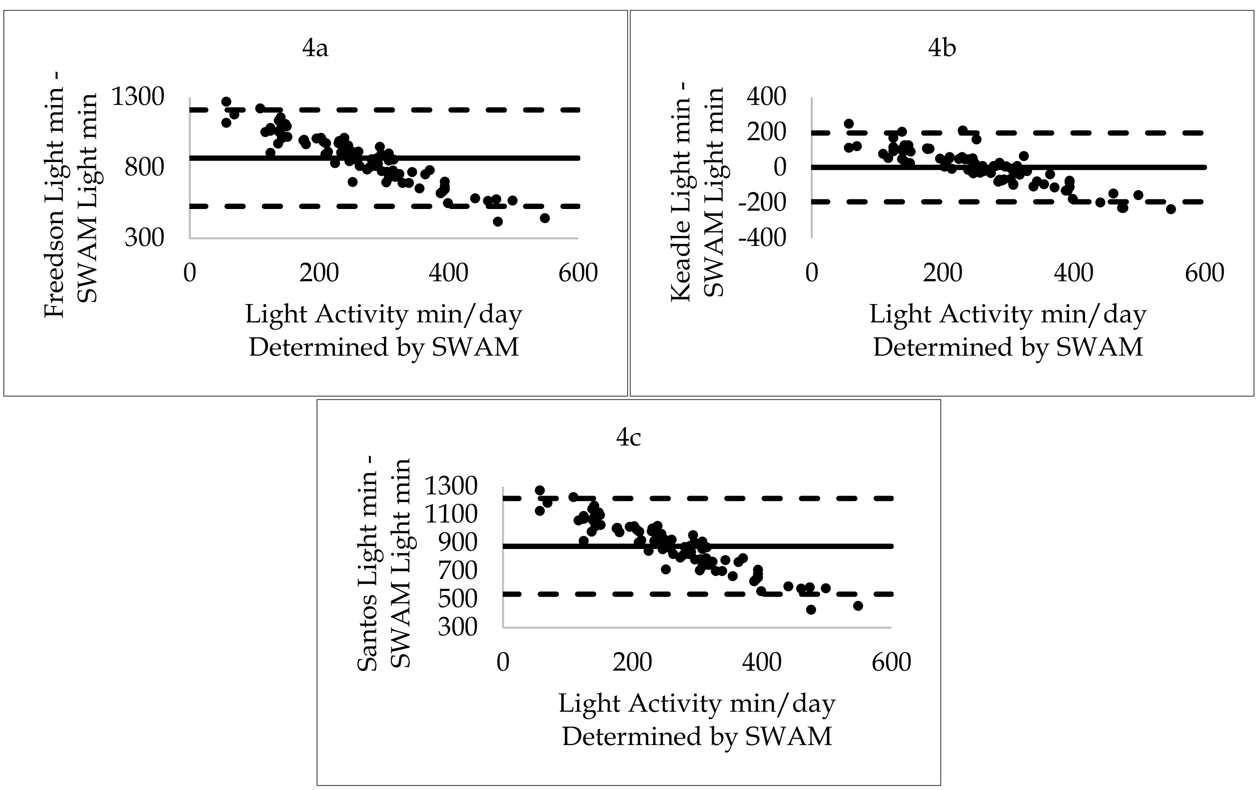

Figure 4. Bland-Altman plots for agreement of time spent in light activity, from three different cut points and the SenseWear Armband Mini monitor (SWAM). 4a) Difference in time spent in light activity between Freedson Adult VM3 2011 cut points (Freedson) and SWAM. 4b) Difference in time spent in light activity between Keadle Women's Health VM 2014 cut points (Keadle) and SWAM. 4c) Difference in time spent in light activity between Santos-Lozano Older Adults 2013 cut points (Santos) and SWAM.

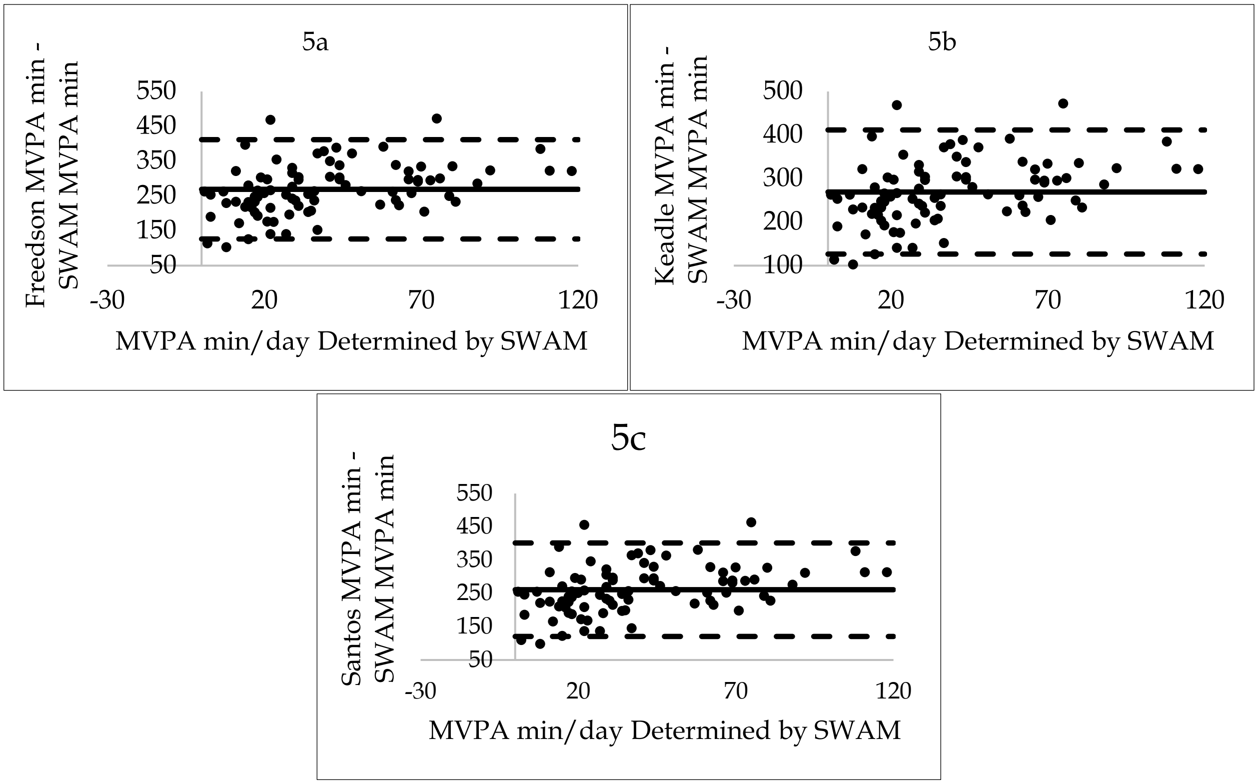

Figure 5. Bland-Altman plots for agreement between time spent in moderate to vigorous physical activity (MVPA) from three different cut points and the SenseWear Armband Mini monitor (SWAM). 5a) Difference in time spent in moderate to vigorous physical activity (MVPA) between Freedson Adult VM3 2011 cut points (Freedson) and SWAM. 5b) Difference in time spent in MVPA between Keadle Women's Health VM 2014 cut points (Keadle) and SWAM. 5c) Difference in time spent in MVPA between Santos-Lozano Older Adults 2013 cut points (Santos) and SWAM.

Associations between CPM from the GT3X+ and PAEE determined by the SWAM and DLW and indirect calorimetry were 0.608 and 0.625, respectively. CCs between CPM from the GT3X+ and PAEE determined by the SWAM and DLW and Indirect Calorimetry were 0.303 and 0.349 respectively.

As shown in Table 6, the mean overall rate of agreement between the two devices was 52.07 ± 6.4%. MVPA showed the highest rate of agreement of all the activity intensities with a mean agreement of 85.36 ± 12.29%. Light intensity activity showed the poorest agreement rates with a mean of 25.19 ± 7.90%. When the GT3X+ did not correctly identify light intensity activity, 71.96 ± 10.25% of the epochs were classified as MVPA. The low agreement rates during light intensity activity and the misclassification of light intensity activity as MVPA indicate an overestimation of activity intensity of the GT3X+ when compared to the SWAM.

Table 6. Epoch by Epoch Analysis of Agreement with SWAM (criterion)

	Mean	SD	CI
Overall agreement	52.07%	6.40%	(50.68% - 53.46%)
Sedentary agreement	57.07%	7.07%	(55.54% - 58.60%)
Sedentary classified as light	32.40%	4.55%	(31.41% - 33.39%)
Sedentary classified as MVPA	10.53%	4.43%	(9.57% - 11.49%)
Light agreement	25.19%	7.90%	(23.48% - 26.91%)
Light classified as sedentary	2.84%	3.53%	(2.08% - 3.61%)
Light classified as MVPA	71.96%	10.25%	(69.74% - 74.19%)
MVPA agreement	85.36%	12.29%	(82.69% - 88.03%)
MVPA classified as sedentary	0.84%	2.01%	(0.41% - 1.28%)
MVPA classified as light	13.80%	11.44%	(11.31% - 16.28%)

Agreement is defined as incidences where the GT3X+ classified activity intensity correctly when compared the SWAM (criterion).

DISCUSSION

To our knowledge, this study is the first to examine the performance of the wrist-worn GT3X+ in measuring PAEE and time in activity intensity in a healthy non-obese population of older women. Although the GT3X+ did not perform well when compared to the SWAM and DLW and indirect calorimetry, our results are useful in informing future research and use of the GT3X+ accelerometer.

The two equations used to estimate PAEE from the GT3X+ were the VM3 and VM3 combo equations. These equations provided by ActiLife were developed for measuring PAEE from the hip. Our results indicate they did not perform well when used to calculate PAEE when the device was worn on the wrist. An equation that was developed for the wrist-worn GT3X+ may produce more accurate estimates of PAEE. Furthermore, the equations used were not developed for our specific population, indicating a need for more population specific equations for estimating PAEE.

The Keadle cut points were chosen for the Epoch-by-epoch analysis because they showed the correct direction in correlation with the criterion measures and were the only cut points to offer estimations of sedentary time. These estimates correlated well with the sedentary time estimated by the SWAM, but the Bland-Altman plot shows that the Keadle cut point consistently underestimates time spent in sedentary time. In addition to this, the large limits of agreement also show poor agreement. These finding are contrary to those found by Ellis et al (9). Their study compared accuracy of behavior classifications made between the hip and wrist placement of the GT3X+. For seven days, participants wore two accelerometers (one on the hip and one on the wrist) and a camera that captured images every 20 seconds to attain information about true participant behavior. Chi-square tests were used to determine significant difference from true behavior. Both placements significantly overestimated time standing but estimated time sitting and riding in a vehicle were not significantly different from true behavior. Bland-Altman plots for minutes walking showed no bias with increasing time for both the hip and wrist placement, indicating good agreement. The authors concluded that both placements provided accurate estimates of sedentary and walking minutes (9). A reason for our different finding may be the device we chose as our criterion. We compared estimates of time in activity intensity to those from the SWAM while Ellis et al (9) compared their estimates to direct observation of activity. Another reason for our different findings could be the demographics of the participants. Our study included non-obese women while the study by Ellis et al (9) included only overweight and obese women. Furthermore, our study examined activities in a free-living environment which is more complex than the limited behaviors that were observed by Ellis et al (9).

Because the Keadle cut points classified sedentary time like the SWAM did, it was able to produce better estimates of light intensity activity than the rest of the cut points. The correlation and linear regression analyses showed that all cut points were not strongly associated with the SWAM when measuring light intensity activity. However, the Keadle cut points show the smallest limits of agreement and lowest mean bias by the Bland-Altman plots (Figure 4) compared to the other cut points, indicating the best agreement with the SWAM on light intensity activity. The Keadle cut points may have showed the best agreement with the SWAM because it classified sedentary time separately while the other cut points did not.

All cut points produced similar estimates of MVPA. Spearman's correlations for MVPA between all cut points and the SWAM were moderate, while regression coefficients were close to 2, indicating the GT3X+ was overestimating time spent in MVPA at a rate of 2 to 1 when compared to the SWAM. ICCS showed poor reliability and Bland-Altman plots in Figure 5 showed that all cut points overestimate time spent in MVPA and the overestimation increases at higher levels of intensity. Furthermore, the plots show a large mean bias and limits of agreement. All of this indicates poor reliability between the two devices. An overestimation of time spent in MVPA may be explained by an overestimation of PAEE which has been seen in other studies aiming to validate the GT3X+ (17).

A study that saw an overestimation of PAEE at higher intensities was one conducted by McMinn et al (17) who aimed to determine the validity of GT3X+'s ability to estimate PAEE under controlled walking conditions and examined the agreement between the hip and wrist placements. A total of 19 participants, aged 19-53 years old, completed three walking trials: a slow-walking trial, a medium-walking trial, and a fast-walking trial. The study compared PAEE estimates from the GT3X+ to indirect calorimetry (IC) and found that PAEE estimates from the GT3X+ highly correlated with indirect calorimetry (hip: r = 0.82, wrist: r = 0.72). However, Bland-Altman plots revealed that the GT3X+ (hip and wrist) significantly underestimated PAEE during slow walking (mean difference of 0.77 and 1.22 for hip and wrist respectively) and significantly overestimated PAEE during fast walking (mean difference of -1.9 and -0.96 for hip and wrist respectively). No differences were seen between the hip and wrist placements during the medium walking trial. McMinn ultimately concluded that the GT3X+ shows high correlation with indirect calorimetry measured PAEE but poor agreement during slow and fast walking trials, and when worn on the wrist, the GT3X+ tended to underestimate EE at rates above 4 $kcal \cdot min^{-1}$ (17).

For the epoch-by-epoch analysis, the mean overall agreement between the two monitors was 52.07%, with the highest agreement rate happening during the measurement of MVPA (85.36%). However, when measuring light intensity activity, the GT3X+ performed very poorly when compared to the SWAM (25.19% agreement). Furthermore, the analysis revealed that the GT3X+ tends to overestimate light intensity activity when compared to the SWAM, misclassifying light intensity activity as MVPA 71.96% of the time. This discrepancy may be due to the tendency of the GT3X+ to overestimate activity intensity and underestimate sedentary time. This would also explain the high agreement rate with the SWAM for MVPA. The results of our epoch-by-epoch analysis are consistent with the results of our analysis of time spent in activity intensity. Bland-Altman plots in Figure 5 showed that all GT3X+ cut points overestimate time spent in MVPA which can be explained by an overestimation of PAEE that has been seen in the previously mentioned studies (4, 17). Agreement between the two devices on sedentary time was similar to the overall agreement, with an average of 57.07% of the epochs between the two devices agreeing when the SWAM classified activity as sedentary.

Counts per minute from the GT3X+ were compared to PAEE estimates from the SWAM and DLW and IC. Spearman correlations between CPM from the GT3X+ and criterion measures were high. These results indicate that a more accurate estimation of PAEE from the wrist-worn GT3X+ is possible and the problem may lie in the estimation equations. Results of other studies looking at CPM suggest this as well. One study found a Pearson's correlation of 0.88 between CPM from the wrist-worn GT3X and the hip worn GT3X (6). If wrist-worn CPM are strongly correlated with hip worn CPM, it should be possible to get similar estimates of PAEE from both placements. This is significant because the hip placement of accelerometers has been previously validated (1, 5, 11, 20).

Findings by Shiroma et al (27) can help to explain the overestimation of PAEE and activity intensity that we saw in our study. Shiroma et al (27) examined how wear placement of the GT3X+ would affect accelerometer output. The GT3X+ was worn on each wrist along with one on the hip for seven days. During this time, the devices worn on the wrist produced significantly higher CPM than the device worn on the hip (27). Since CPM are used to calculate PAEE and activity intensity, and the equations are developed for the hip placement, a higher CPM from the wrist placement would explain the overestimation of PAEE and activity intensity.

A study by Kamanda et al (13) explored 11 different cut points ranging from < 500 to < 10,000 CPM for sedentary behavior and from ≥ 3000 to ≥ 15,000 CPM for MVPA on data from a wrist worn GT3X+. Cut points were also applied to concurrent data from a hip worn GT3X+. The cut points used for the hip worn data were < 200 CPM for sedentary activity and ≥ 2690 CPM for MVPA. Mean difference was assessed between the 11 different wrist cut points and the hip cut points. For sedentary behavior, a cut point of < 2000 CPM showed the smallest mean difference compared to the hip cut points. For MVPA, cut points of ≥ 7500 and ≥ 8250 CPM showed the smallest mean difference compared to the hip cut points. This study shows how vastly different CPM outputs from the hip and wrist placements of the same device can be (13). These results help to explain the overestimation of PAEE and activity intensity we saw from the wrist-worn GT3X+ in our study.

Strengths of this study include the use of DLW and lab measured REE as a criterion measure since it is considered the gold standard in techniques measuring EE. Other strengths include objectively determined PA and sedentary time, long wear time of the activity monitors each day, the number of days the women wore the activity monitors, and matching monitor wear time and DLW measurement period. Furthermore, PA and sedentary time were measured in a free-living environment (33).

Although this study has many strengths, several limitations should be considered. REE was only measured once and was used as a representation of the individuals average REE. Most importantly, the participants enrolled in the study were non-obese inactive older women and it is well known that the validity of accelerometry algorithms may be affected by age and activity level. Therefore, the results of this study cannot be generalized to populations outside of this study sample.

In conclusion, the currently available PAEE estimation equations do not allow us to accurately measure PAEE with the GT3X+ worn on the wrist in a population of older women. Specifically, when compared to the SWAM, the cut points available for the GT3X+ tend to overestimate time spent in light intensity activity and MVPA and underestimate sedentary time when worn on the wrist. This is most likely because there are few PAEE equations and intensity cut points developed for the device when worn on the wrist and none that fit our participant demographic. The high correlation between CPM from the GT3X+ and PAEE determined by criterion measures suggests the potential of developing better equations using the wrist-worn GT3X+ to estimate PAEE and time by activity intensity. Future research should explore the validity of the wrist-worn GT3X+ in different populations and work to develop EE equations for the wrist placement.

ACKNOWLEDGEMENTS

Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Award Number R00AG031297. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest. Results of the present study do not constitute endorsement by the American College of Sports Medicine. Results are stated clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

REFERENCES

1. Bai Y, Welk GJ, Nam YH, Lee JA, Lee J-M, Kim Y, Meier NF, Dixon PM. Comparison of consumer and research monitors under semistructured settings. Med Sci Sports Exerc 48(1): 151-158, 2016.

2. Bergmann JH, Graham S, Howard N, McGregor A. Comparison of median frequency between traditional and functional sensor placements during activity monitoring. Measurement (Lond) 46(7): 2193-2200, 2013.

3. Breneman CB, Kline CE, West D, Sui X, Wang X. The effect of structured exercise on sleep during the corresponding night among older women in an exercise program. Journal of Aging and Physical Activity 27(4): 482-488, 2019.

4. Ceaser TG. The estimation of caloric expenditure using three triaxial accelerometers. [Dissertation]: University of Tennesee; 2012.

5. Chen KY, Acra SA, Majchrzak K, Donahue CL, Baker L, Clemens L, Sun M, Buchowski MS. Predicting energy expenditure of physical activity using hip-and wrist-worn accelerometers. Diabetes Technol The 5(6): 1023-1033, 2003.

6. Dieu O, Mikulovic J, Fardy PS, Bui-Xuan G, Béghin L, Vanhelst J. Physical activity using wrist-worn accelerometers: comparison of dominant and non-dominant wrist. Clin Physiol Funct Imaging 37(5): 525-529, 2017.

7. Dipietro L, Caspersen CJ, Ostfeld AM, Nadel ER. A survey for assessing physical activity among older adults. Med Sci Sports Exerc 25(5): 628-642, 1993.

8. Ekelund U, Franks PW, Sharp S, Brage S, Wareham NJ. Increase in physical activity energy expenditure is associated with reduced metabolic risk independent of change in fatness and fitness. Diabetes Care 30(8): 2101-2106, 2007.

9. Ellis K, Kerr J, Godbole S, Staudenmayer J, Lanckriet G. Hip and wrist accelerometer algorithms for free-living behavior classification. Med Sci Sports Exerc 48(5): 933, 2016.

10. Gastin P, Cayzer C, Robertson S, Dwyer D. Validity of the actigraph gt3x+ and sensewear armband to predict energy expenditure during physical activity and sport. J Sci Med Sport 20: e103, 2017.

11. Godfrey A, Conway R, Meagher D, ÓLaighin G. Direct measurement of human movement by accelerometry. Med Eng Phys 30(10): 1364-1386, 2008.

12. Hildebrand M, VT VH, Hansen BH, Ekelund U. Age group comparability of raw accelerometer output from wrist-and hip-worn monitors. Med Sci Sports Exerc 46(9): 1816-1824, 2014.

13. Kamada M, Shiroma EJ, Harris TB, Lee I-M. Comparison of physical activity assessed using hip-and wrist-worn accelerometers. Gait & posture 44: 23-28, 2016.

14. Keadle SK, Shiroma EJ, Freedson PS, Lee I-M. Impact of accelerometer data processing decisions on the sample size, wear time and physical activity level of a large cohort study. BMC Public Health 14(1): 1-8, 2014.

15. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, Nathan DM. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Eng J Med 346(6): 393-403, 2002.

16. Manini T, Everhart J, Patel K, Schoeller D, Colbert L, Visser M, Tylavsky F, Bauer D, Goodpaster B, Harris T. Daily activity energy expenditure and mortality among older adults. JAMA 296(2): 171-179, 2006.

17. McMinn D, Acharya R, Rowe DA, Gray SR, Allan JL. Measuring activity energy expenditure: accuracy of the gt3x+ and actiheart monitors. Int J Exerc Sci 6(3): 5, 2013.

18. Navalta JW, Stone WJ, Lyons TS. Ethical issues relating to scientific discovery in exercise science. International journal of exercise science 12(1): 1, 2019.

19. Noury N, Galay A, Pasquier J, Ballussaud M. Preliminary investigation into the use of autonomous fall detectors. In Proceedings of the Annu Int Conf IEEE Eng Med Biol Soc. 2008. p. 2828-2831.

20. Plasqui G, Bonomi AG, Westerterp KR. Daily physical activity assessment with accelerometers: new insights and validation studies. Obes Rev 14(6): 451-462, 2013.

21. Reed GW, Hill JO. Measuring the thermic effect of food. Am J Clin Nutr 63(2): 164-169, 1996.

22. Sagelv EH, Hopstock LA, Johansson J, Hansen BH, Brage S, Horsch A, Ekelund U, Morseth B. Criterion validity of two physical activity and one sedentary time questionnaire against accelerometry in a large cohort of adults and older adults. BMJ Open SEM 6(1): e000661, 2020.

23. Santos-Lozano A, Santin-Medeiros F, Cardon G, Torres-Luque G, Bailon R, Bergmeir C, Ruiz JR, Lucia A, Garatachea N. Actigraph gt3x: validation and determination of physical activity intensity cut points. Int J Sports Med 34(11): 975-982, 2013.

24. Sasaki JE, John D, Freedson PS. Validation and comparison of actigraph activity monitors. J Sci Med Sport 14(5): 411-416, 2011.

25. Schoeller DA. Recent advances from application of doubly labeled water to measurement of human energy expenditure. J Nutr 129(10): 1765-1768, 1999.

26. Sherman SE, D'Agostino RB, Cobb JL, Kannel WB. Does exercise reduce mortality rates in the elderly? Experience from the Framingham heart study. Am Heart J 128(5): 965-972, 1994.

27. Shiroma EJ, Schepps MA, Harezlak J, Chen KY, Matthews CE, Koster A, Caserotti P, Glynn NW, Harris TB. Daily physical activity patterns from hip- and wrist-worn accelerometers. Physiol Meas 37(10): 1852-1861, 2016.

28. Stob NR, Bell C, van Baak MA, Seals DR. Thermic effect of food and β-adrenergic thermogenic responsiveness in habitually exercising and sedentary healthy adult humans. J Appl Physiol 103(2): 616-622, 2007.

29. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. Br J Sports Med 48(13): 1019-1023, 2014.

30. Tuomilehto J, Lindström J, Eriksson JG, Valle TT, Hämäläinen H, Ilanne-Parikka P, Keinänen-Kiukaanniemi S, Laakso M, Louheranta A, Rastas M. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N Eng J Med 344(18): 1343-1350, 2001.

31. van Hees VT, Renstrom F, Wright A, Gradmark A, Catt M, Chen KY, Lof M, Bluck L, Pomeroy J, Wareham NJ. Estimation of daily energy expenditure in pregnant and non-pregnant women using a wrist-worn tri-axial accelerometer. PLoS ONE 6(7)2011.

32. Wang X, Bowyer KP, Porter RR, Breneman CB, Custer SS. Energy expenditure responses to exercise training in older women. Physiol 5(15): e13360, 2017.

33. Wang X, Breneman CB, Sparks JR, Blair SN. Sedentary time and physical activity in older women undergoing exercise training. Med Sci Sports Exerc 52(12): 2590-2598, 2020.