Gait & Posture 68 (2019) 141–149

Contents lists available at ScienceDirect

Gait & Posture

journal homepage: www.elsevier.com/locate/gaitpost

Full length article

Calibration of wrist-worn ActiWatch 2 and ActiGraph wGT3X for assessment of physical activity in young adults

Paul Lee[a,*], C.Y. Tse[b]

[a] School of Nursing, Hong Kong Polytechnic University, Hong Kong
[b] Department of Health and Physical Education, Education University of Hong Kong, Hong Kong

ARTICLE INFO

Keywords:
Accelerometry
Calibration
Cut-points
Measurement
Motion sensor
Physical activity

ABSTRACT

Background: The validity of Actiwatch 2 in assessing sleep was evident, but its validity in assessing physical activity (PA) level was unknown.

Research question: The objective of this study was to validate the wrist-worn Actiwatch 2 and ActiGraph wGT3X as a measurement of PA level against energy expenditure measured by indirect calorimetry.

Methods: Twenty-seven university students aged 18–26 were recruited from July 2016 to May 2017. They were instructed to run at different speeds (4, 6, 8, 10, and 12 km/h) on a treadmill, each speed for 10 min. Oxygen consumption and carbon dioxide production of the subjects was measured by indirect calorimetry using the Cosmed K4b^2 gas analyzer. Each subjects wore a single pair of accelerometers (Actiwatch 2 and ActiGraph wGT3X) on both wrists.

Results: All the accelerometers were strongly correlated ($\rho = 0.83$-0.94, all p-values < 0.001), and all four accelerometers were strongly correlated with the METs obtained from the Cosmed K4b^2 ($\rho = 0.72$-0.74, all p-values < 0.001). Regression analysis showed that the non-dominant wrist-worn Actiwatch 2 cutoff cpm for moderate and vigorous PA were 399 and 1,404, respectively; for the ActiGraph wGT3X-BT the corresponding cutoffs were 4,514 and 15,044, respectively. The goodness-of-fit of the MET prediction equations were all > 75%. When classifying the activities as either sedentary, light activity, moderate-intensity activity, or vigorous-intensity activity using the MET prediction equations, the agreements between the four accelerometers and that by the Cosmed K4b^2 were high, all AUCs were above 80% except those of the Actiwatch worn on the left (non-dominant) wrist. The Bland-Altman plots show that, for all four accelerometers, the biases were close to zero and error variances were largest when the mean measurements were around 6 METs.

Significance: We showed that wrist-worn Actiwatch 2 and ActiGraph wGT3X-BT were strongly correlated in PA assessment.

1. Introduction

An accelerometer is an electronic motion sensor that measures the acceleration, or change of velocity. It can be worn on wrist or waist to measure body movement. The intensity of the body movement measured by accelerometer can proxy the activity level of the wearer, and it can be used to measure both physical activity (PA) level and sleeping pattern. Accelerometers have been validated against doubly-labeled water [1], the gold standard of physical activity measurement, and polysomnography (PSG) [2], the gold standard of sleep measurement. In addition, for young children from whom self-report is infeasible, an accelerometer is a valid method in assessing their PA level [3,4] and sleep pattern [5]. Given the increasingly affordable cost of accelerometers, researchers have shifted the measurement of habitual PA level and sleep pattern from self-reported questionnaires to accelerometers [6].

Currently, in studies that measure both PA and sleep, researchers have to use two types of accelerometer, one for measuring PA and another for measuring sleep [7]. Because of the high cost of some sleep accelerometers (for example, the Actiwatch 2 by Phillips Respironics Mini-Mitter costs about US $1500), it would be logical to produce a single accelerometer that can measure both PA and sleep. We previously demonstrated that the ActiGraph Link accelerometer yielded similar sleeping parameters to the Actiwatch 2 (validated against PSG) in a free-living condition [8], and the evidence supported the ActiGraph Link as a valid measurement of both PA and sleep. The objective of the current study was to validate wrist-worn Actiwatch 2 as a measurement of PA level against energy expenditure measured by indirect

* Corresponding author at: School of Nursing, Hong Kong Polytechnic University, GH527, Hung Hom, Kowloon, Hong Kong.
E-mail address: paul.h.lee@polyu.edu.hk (P. Lee).

https://doi.org/10.1016/j.gaitpost.2018.11.023
Received 19 May 2018; Received in revised form 15 November 2018; Accepted 17 November 2018
0966-6362/

Table 1
Spearman correlation of the Actiwatch count, ActiGraph VM, and Cosmed K4b^2 METs minute-by-minute data (n = 27, count = 2135).

	Actiwatch (left wrist)	ActiGraph VM (right wrist)	ActiGraph VM (left wrist)	Cosmed K4b^2
Actiwatch (right wrist)	0.93 (n = 14, count = 1080)	0.93 (n = 20, count = 1549)	0.89 (n = 20, count = 1549)	0.73 (n = 20, count = 1549)
Actiwatch (left wrist)		0.83 (n = 18, count = 1386)	0.85 (n = 18, count = 1386)	0.72 (n = 18, count = 1386)
ActiGraph VM (right wrist)			0.94	0.73
ActiGraph VM (left wrist)				0.74

MET: metabolic equivalent tasks; VM: vector magnitude.
All correlations were significant at $p < 0.001$.

calorimetry. As a secondary objective, the subjects also wore ActiGraph wGT3X, and its associations with Actiwatch 2 and indirect calorimetry were assessed. We hypothesized that the energy expenditure measured by Actiwatch 2, ActiGraph GT3X, and indirect calorimetry (Cosmed K4b^2 gas analyzer) should have strong agreement in measuring physical activity.

2. Methods

2.1. Participants

Twenty-seven university students aged 18–26 were recruited from July 2016 to May 2017. Those who could not perform the necessary PAs and pregnant women were excluded from this study. An information sheet was provided to all subjects prior to the study and written consent was obtained. They were informed of the right to withdraw during any part of the study. Each subject was given a coupon for HK $100 upon completion of the study. The study protocol was approved by the Human Subjects Ethics Sub-committee of the Hong Kong Polytechnic University (code: HSEARS20140922001).

2.2. Procedure

The test was conducted in a laboratory with air conditioning. The body weight and height of the subjects were measured with an electronic scale and stadiometer (FTS-H fitness scale, DPS Promatic Srl, Forlì, Italy) before the test. Each subject was instructed to lie down before the test for at least 20 min to measure their resting energy expenditure. They were instructed to run at different speeds (4, 6, 8, 10, and 12 km/h) on a treadmill, each speed for 10 min. Running of these speeds correspond to METs of 3.0, 6.0, 9.0, 10.5, and 11.8, respectively [9]. They were given a 5-minute recovery period between each run (corresponds to METs of 1.5). During the test, each subject's heart rate was monitored in real-time. The test would be terminated if the heart rate reached 90% of their maximum heart rate.

2.3. Energy expenditure

Oxygen consumption and carbon dioxide production of the subjects was measured by indirect calorimetry using the Cosmed K4b^2 gas analyzer [10]. It was warmed-up for a minimum of 60 min before each test. Energy expenditure of the subjects in a 1-minute epoch was calculated according to the Weir equation. Metabolic equivalent tasks (METs) of the subjects in each 1-minute epoch were calculated by dividing the energy expenditure in the epoch by the resting energy expenditure. We used a personalized estimate of resting energy expenditure instead of the standard formula of 3.5 $\text{kcal kg}^{-1} \text{min}^{-1}$, because many people have a resting energy expenditure lower than 3.5 [11].

2.4. Accelerometers

Each subjects wore four accelerometers, two on each wrist. That is, they wore Actiwatch 2 and ActiGraph wGT3X accelerometers on both the left and right wrists. Due to the specific design of the Actiwatch 2 (that it fits at wrist but not at waist), we decided not to put any accelerometers at the subjects' waist. The vector magnitude of the ActiGraph, computed according to the Pythagoras theorem, was used to summarize the count data of the three dimensions. All accelerometers and the Cosmed K4b^2 were synchronized with the same computer, so that their built-in clocks were synchronized. The epoch length was set at 1 min.

2.5. Statistical analysis

We chose pooled male and female subjects because similar calibration studies have shown there are no sex differences [12]. Spearman correlation was used to examine the associations between the count-per-minute (cpm) data of all accelerometers and the METs obtained from the Cosmed K4b^2. Regression was used to establish the predictive relationship between the accelerometer cpm and METs. Spearman correlation was used as we expected the cpm and METs were distributed as a mixture of uniform distribution across 3.0, 6.0, 9.0, 10.5, and 11.8 METs (running tasks) and normally distribution of mean 1.5 (resting period). In addition to a linear term of accelerometer cpm, a quadratic term was also included to approximate the monotonic increasing and decreasing slope trend between accelerometer cpm and METs. Log association was also tested, but quadratic association was selected because it showed better goodness-of-fit (in terms of R^2) in our data. Multilevel regression with random intercept and slope was used, because it allowed each subject to have his or her own intercept and slope values. Residual plots showed that the residuals for all four regressions were normally distributed. The averages of the individual intercepts and slopes were used as the MET prediction equation. The cpm corresponding to the cutoffs for light activity, moderate-intensity activity, and vigorous-intensity activity were obtained by equating the regression equation to 1.5, 3.0, and 6.0, respectively. Since the association between accelerometer cpm and METs should be monotonic increasing (i.e., an increase of accelerometer cpm should represent a larger amount of energy expenditure) and we used a quadratic equation model this association, the equation beyond the maximum point of the quadratic equation is not applicable in modeling the association between accelerometer cpm and METs and should be removed. This leads to the removal of the larger root of the quadratic equation, therefore the smaller of the two roots of the quadratic equation was used. R^2 was be used to assess the goodness-of-fit of the regression models.

Using the cutoffs obtained from the above regression models, the accelerometer cpm were transformed to METs and further transformed to level of physical activity (< 1.5 METs: sedentary, 1.5–2.99: light, 3–5.99: moderate-intensity; ≥6.0: vigorous intensity). The agreements between activity level classifications, including the agreement between each accelerometer with the Cosmed K4b^2, agreement between the two accelerometers worn on the same wrist, and agreement between the same brand of accelerometer worn on different wrists, were presented using the confusion matrix. Furthermore, these agreements were evaluated using the Bland-Altman plot, sensitivity, specificity, and area

Table 2
Results of the multilevel regression predicting MET using accelerometer counts and there estimated cutoff points for different levels of activities.

	MET prediction equation	R^2	Cutoff		
			Light activity or above (≥1.5 METs)	Moderate-intensity activity or above (≥3 METs)	Vigorous-intensity activity (≥6 METs)
Actiwatch (right wrist)	$1.26 + 0.00403 * \text{count} - 3.31 * 10^{-7} * \text{count}^2$	75.9%	61	450	1322
Actiwatch (left wrist)	$1.22 + 0.00488 * \text{count} - 1.05 * 10^{-6} * \text{count}^2$	78.6%	58	399	1404
ActiGraph VM (right wrist)	$1.35 + 0.000366 * \text{count} - 4.43 * 10^{-9} * \text{count}^2$	81.6%	418	4793	15696
ActiGraph VM (left wrist)	$1.41 + 0.000371 * \text{count} - 4.42 * 10^{-9} * \text{count}^2$	80.9%	232	4514	15044

MET: Metabolic equivalent task; VM: vector magnitude.

under the receiver operating characteristic (ROC) curve (AUC). The sensitivity, specificity, and AUC for classifying light activity or above (≥1.5 METs), moderate-intensity activity or above (≥3 METs), and vigorous-intensity activity (≥6 METs) were computed. For same-model comparison, the accelerometer worn in the left wrist was used as the reference. For same-wrist comparison, the ActiGraph was used as the reference. All analyses were conducted using R 3.3.0 (https://www.r-project.org) and multilevel regression was fitted using package *lme4* [13].

3. Results

The sex of the sample was balanced with 12 males (44.4%) and 15 females. Their mean height, weight, and BMI were 168.7 cm (SD 8.3 cm), 62.4 kg (SD 10.9 kg), and 21.9 (SD 3.2), respectively. All subjects were right-handed via self-report. During the tests, the two Actiwatches malfunctioned several times and all available data included in the analysis are shown in Table 1. In sum, the 27 subjects provided 2,135 min-by-minute accelerometer and indirect calorimetry data. Table 1 shows that all the accelerometer cpm were strongly correlated with Spearman correlations ranging from 0.83 to 0.94 (all p values were < 0.001), and all four accelerometers were strongly correlated with the METs obtained from the Cosmed K4b^2 with Spearman correlations ranging from 0.72 to 0.74 (all p values were < 0.001).

Table 2 shows the regression result and the estimated cutoff points for different levels of activity. Fig. 1 shows the scatter plots and regression equation plots of accelerometer cpm and METs by indirect calorimetry. For Actiwatch 2 left wrist, the cpm beyond maximum METs of the prediction equation ($0.00488 / 2 * 1.05 * 10^{-6} = 2324$) was discarded. The goodness-of-fit of the MET prediction equations for ActiGraph (80.9% and 81.6%) were higher than that of the Actiwatch (75.9% and 78.6%). For both accelerometer brands, the cutoff points for all levels of activity for the left wrist (i.e., the non-dominant wrist) were slightly lower. There were substantial differences between the prediction equation for right and left Actiwatch, and this difference was caused by the influential observations for the right Actiwatch with cpm > 3000 and METs < 6.

According to the cutoff points estimated using multilevel regression, accelerometer cpm for the Actiwatch and ActiGraph were classified as either sedentary, light activity, moderate-intensity activity, or vigorous-intensity activity. The agreement of the activity level classification between the four accelerometers and that by the Cosmed K4b^2 is shown in Table 3. All AUCs were above 80% except those of the Actiwatch worn on the left (non-dominant) wrist. All four accelerometers had high sensitivities (> 90%) in classifying light activity or above (≥1.5 METs), but the sensitivities in classifying vigorous-intensity activity (≥6.0 METs) were about 60% only, except for those of the Actiwatch worn on the left (non-dominant) wrist. The Bland-Altman plots (Fig. 2) show that, for all four accelerometers, the biases were close to zero and error variances were largest when the mean measurements were around 6 METs. The limits of agreement were around 4 METs in general for Actiwatch (right wrist: 4.17, left wrist: 3.89) and were higher than that of ActiGraph (right wrist: 3.68, left wrist: 3.75). The ROC curves for all the above classifications are shown in Supplementary Fig. 1.

Table 4 shows the agreement of the activity level classification between the four accelerometers. Same-model accelerometers worn on different wrist had strong agreement with sensitivities greater than 84%. Same-wrist comparison showed that Actiwatch and ActiGraph worn on the left (non-dominant) wrist had relatively weaker agreement in classifying vigorous-intensity activity. The Bland-Altman plots (Fig. 3) show that the biases between the same accelerometer model worn on different wrist and same-wrist comparison of the two brands were close to zero. The limit of agreement of ActiGraph was smaller than that of Actiwatch. Furthermore, the agreement between the Actiwatches was weak when METs were high, and the error variances were largest when the mean measurements were around 6 METs for same-

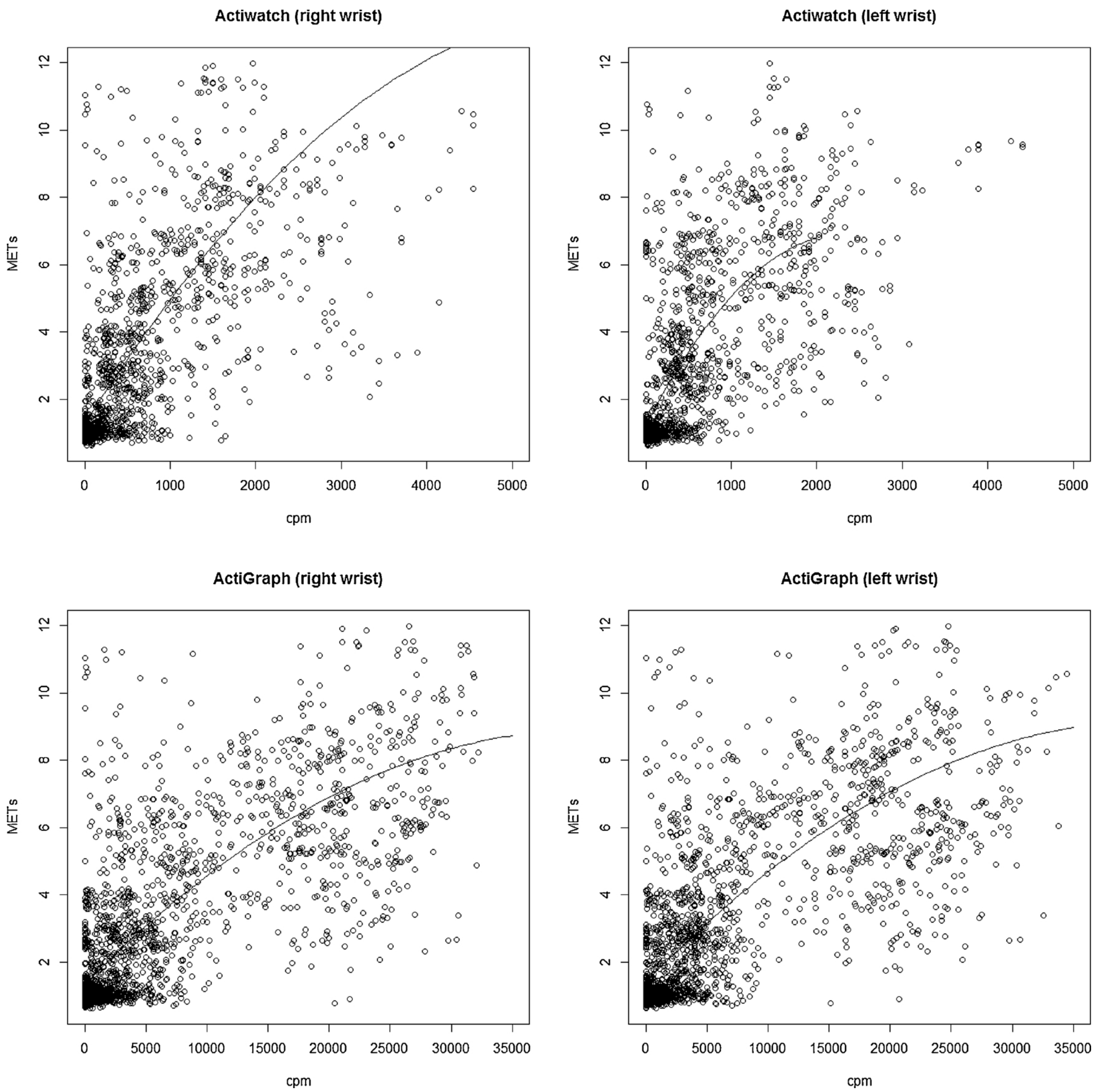

Fig. 1. Scatter plots and regression equation plots of accelerometer count per minute (cpm) and metabolic equivalent tasks (METs) obtained using Cosmed K4b^2 gas analyzer.

wrist accelerometers. It appeared that the right Actiwatch constantly overestimated METs over that of the left Actiwatch, as confirmed by a regression analysis ($MET_{LeftActiwatch} = 1.07 + 0.59MET_{RightActiwatch}$).

4. Discussion

To the authors' best knowledge, this is the first calibration study of Actiwatch against indirect calorimetry. Our results showed that Actiwatch was strongly correlated with ActiGraph and these two accelerometers had similar criterion validity against indirect calorimetry. Actiwatch was designed for sleep measurement and has been validated against PSG [2]. Together with our results, the evidence supports the use of Actiwatch 2 in measuring 24-hour human activity. Note that the accuracy of both accelerometers was low for high physical activity level, in particular for METs > 8. However, as long as these accelerometers can correctly distinguish moderate-to-vigorous intensity physical activity with light activity and sedentary behavior, the accelerometers can be used to measure human activity. Nonetheless, non-dominant-wrist worn Actiwatch 2 and ActiGraph GT3X had limited agreement in vigorous-intensity activity, and such a difference should be acknowledged when using either accelerometer to measure vigorous-intensity activity.

Our calibration study on wrist-worn ActiGraph was also unique in the sense that other existing studies did not use 1-min epoch (e.g., 1-sec [14] or 15-sec [15] epoch were used). As wrist-worn accelerometers demonstrated lower criterion validity than waist-worn accelerometers [14], we suspected that higher resolution data, that is, movement data in smaller epoch length, is required to achieve an acceptable criterion validity for wrist-worn accelerometers. However, with current technological limitations, the battery life could not support habitual 24-hour activity measurement with high resolution data. For example, with 1-minute epoch, the Actiwatch 2 can support recording time of 30.3 days. This implies that if a standard 7-day accelerometer measurement period is to be used, the minimum epoch length would be 15-sec (30.3/4 = 7.6 days), and if 1-second epoch is chosen (although this option is not available in the Actiwatch communication software), the battery could only support measurement for 30.3/60 days (approximately half a day). Given the battery constraint, our results of ActiGraph 1-min epoch data calibration is necessary and applicable to current research.

In our homogeneous sample with minimal confounding effects, we found that the cutoff points of physical activity classification of ActiGraph obtained in this study were very different from studies calibrating waist-worn accelerometers. Regression analysis showed that the cutoff cpm for moderate and vigorous PA were 4514 and 15,044,

Table 3
Agreement of the activity level classification by Actiwatch count and ActiGraph VM with that by Cosmed K4b^2 METs (n = 27, count = 2135).

		Sedentary (< 1.5 METs)	Cosmed K4b^2				Sensitivity	Specificity	AUC
			Light activity (1.5-2.99 METs)	Moderate-intensity activity (3-5.99 METs)	Vigorous-intensity activity (≥6 METs)				
Actiwatch (right wrist)	Sedentary (< 1.5 METs)	377	34	12	9				
	Light activity (1.5-2.99 METs)	241	105	87	34	Light activity or above (≥1.5 METs)	93.7%	55.2%	86.5%
	Moderate-intensity activity (3-5.99 METs)	62	87	145	70	Moderate-intensity activity or above (≥3 METs)	77.3%	81.9%	86.4%
	Vigorous-intensity activity (≥6 METs)	3	15	87	182	Vigorous-intensity activity (≥6 METs)	61.7%	91.6%	84.5%
Actiwatch (left wrist)	Sedentary (< 1.5 METs)	292	10	4	16				
	Light activity (1.5-2.99 METs)	245	60	75	24	Light activity or above (≥1.5 METs)	96.2%	49.1%	88.7%
	Moderate-intensity activity (3-5.99 METs)	58	105	145	127	Moderate-intensity activity or above (≥3 METs)	79.9%	76.5%	83.6%
	Vigorous-intensity activity (≥6 METs)	0	23	79	13	Vigorous-intensity activity (≥6 METs)	42.4%	90.7%	78.5%
ActiGraph VM (right wrist)	Sedentary (< 1.5 METs)	457	57	24	10				
	Light activity (1.5-2.99 METs)	430	178	127	42	Light activity or above (≥1.5 METs)	92.4%	48.9%	85.4%
	Moderate-intensity activity (3-5.99 METs)	45	81	165	109	Moderate-intensity activity or above (≥3 METs)	76.4%	88.0%	87.1%
	Vigorous-intensity activity (≥6 METs)	2	25	130	253	Vigorous-intensity activity (≥6 METs)	61.1%	90.9%	85.7%
ActiGraph VM (left wrist)	Sedentary (< 1.5 METs)	367	42	12	4				
	Light activity (1.5-2.99 METs)	523	191	153	50	Light activity or above (≥1.5 METs)	95.2%	39.3%	84.7%
	Moderate-intensity activity (3-5.99 METs)	42	82	132	108	Moderate-intensity activity or above (≥3 METs)	74.5%	88.1%	86.5%
	Vigorous-intensity activity (≥6 METs)	2	26	149	252	Vigorous-intensity activity (≥6 METs)	60.9%	89.7%	85.3%

AUC: Area under the receiver operating characteristic (ROC) curve; MET: Metabolic equivalent task; VM: vector magnitude.

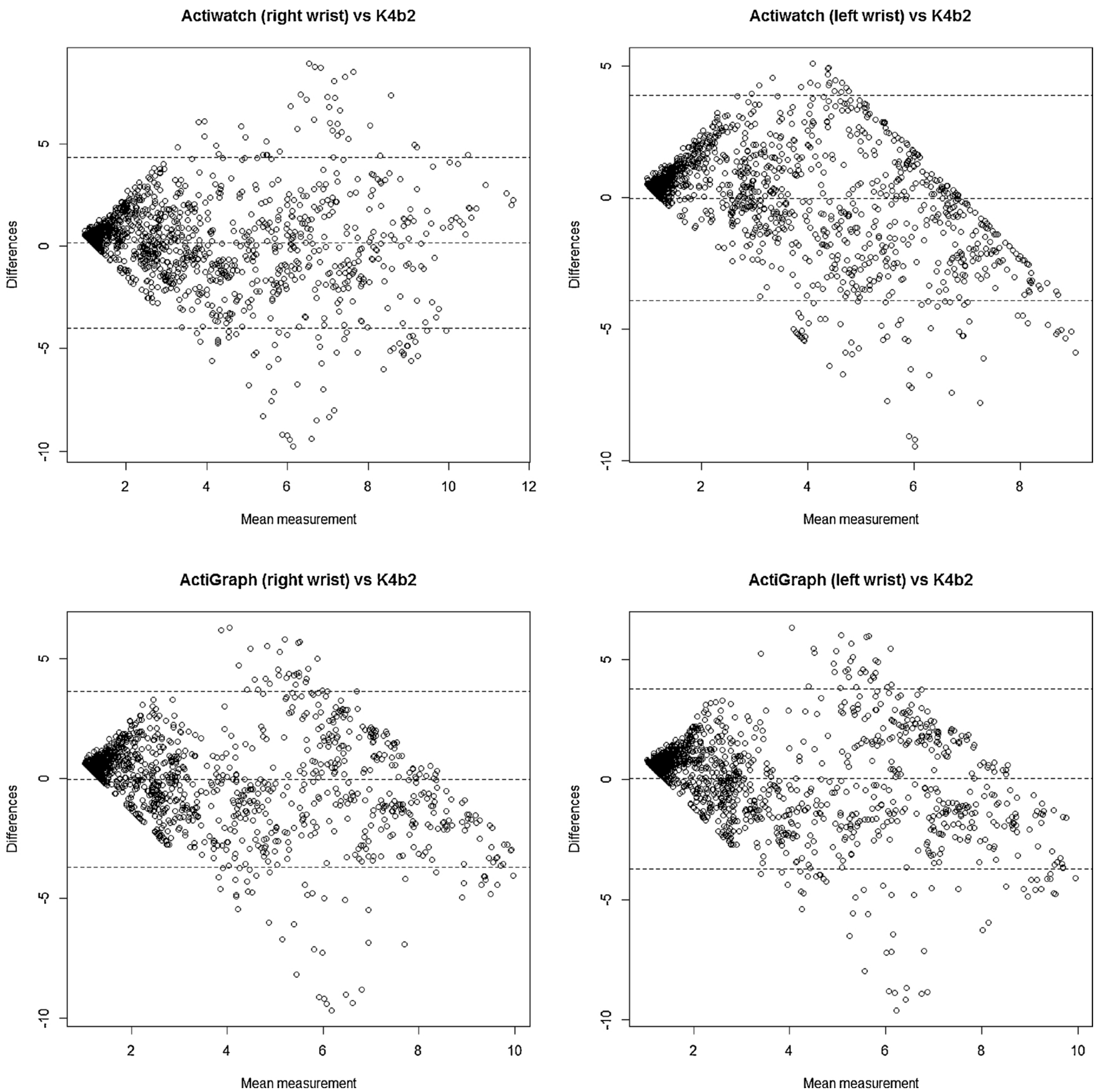

Fig. 2. Bland-Altman plots for Actiwatch 2 and ActiGraph wGT3X-BT with Cosmed $K4b^2$ gas analyzer.

respectively. These cutoff points are very different from those obtained in waist-worn ActiGraph calibration studies among children and adolescents, in which the highest vector magnitude cutoff cpm for moderate and vigorous PA were 3028 and 4,448, respectively, as shown in a recent review [16]. We postulated that during high-speed running, the arm swinging movement was much more dynamic than the waist twisting movement, so the wrist-worn ActiGraph was more sensitive than the waist-worn ActiGraph in classifying vigorous PA.

It was expected that cpm of ActiGraph and Actiwatch were highly correlated as they both measure accelerations. Similar research had also shown that ActiGraph GT3X was highly correlated with other accelerometer brands including GENEActiv [17] and Axivity AX3 [18]. The high but not perfect correlation found across different accelerometer brands could be explained by their difference in raw acceleration data handling methods. Transforming the raw acceleration data to cpm involves frequency filtering, noise attenuating, band-pass filtering, rectification, and integration [18]. If the same transformation protocol was applied onto different brands of accelerometer, the cpm data yielded would be nearly perfectly correlated (Cohen's kappa = 0.95 [18]). However, we could not extract the raw acceleration data from Actiwatch as the computer software provided by the manufacturer did not allow so, therefore we could not examine the correlation of the raw acceleration between Actiwatch and ActiGraph.

The strength of this study lies in the inclusion of the quadratic term in the calibration equation. Although it makes sense to assume a linear association between accelerometer cpm and energy expenditure, our results showed that this association is concave downward. That is, the strength of association between the amount of energy spent and wrist movement reduced with exercise intensity. This is called the *plateau phenomenon* and has been demonstrated in another accelerometer calibration study [19]. Subjects in that study had provided data at a treadmill speed of 12 km/h, but these data were removed as they attenuated the mis-specified linear association. With these data removed, the prediction accuracy of calibration equation on moderate and vigorous intensity PA improved. By modeling such a concave, monotonic increasing association with a quadratic equation, the goodness-of-fit and predictive performance of our proposed calibration model is an improvement over the commonly-used linear model. Empirically, the quadratic models (average R-square of the four regression models = 79.3%) had a better fit to the data than the linear model (average R-square of the four regression models = 76.6%). Therefore, our equation can be used to classify very vigorous intensity PA because the plateau phenomenon was modeled by the quadratic term. In older adults who seldom engage in vigorous PA, the plateau phenomenon may not exist [11] and we can assume a linear association between body movement and energy expenditure.

Table 4
Agreement of the activity level classification by Actiwatch count and ActiGraph VM (n = 27, count = 2135).

		Sedentary (< 1.5 METs)	Actiwatch (left wrist)				Sensitivity	Specificity
			Light activity (1.5-2.99 METs)	Moderate-intensity activity (3-5.99 METs)	Vigorous-intensity activity (≥6 METs)			
Actiwatch (right wrist)	Sedentary (< 1.5 METs)	207	87	1	0			
	Light activity (1.5-2.99 METs)	45	209	75	0	Light activity or above (≥1.5 METs)	89.4%	82.1%
	Moderate-intensity activity (3-5.99 METs)	0	35	200	14	Moderate-intensity activity or above (≥3 METs)	84.6%	93.5%
	Vigorous-intensity activity (≥6 METs)	0	3	73	132	Vigorous-intensity activity (≥6 METs)	90.4%	91.9%
			ActiGraph VM (left wrist)					
ActiGraph VM (right wrist)	Sedentary (< 1.5 METs)	343	204	1	0			
	Light activity (1.5-2.99 METs)	81	641	53	2	Light activity or above (≥1.5 METs)	88.0%	80.7%
	Moderate-intensity activity (3-5.99 METs)	1	72	288	39	Moderate-intensity activity or above (≥3 METs)	92.9%	94.6%
	Vigorous-intensity activity (≥6 METs)	0	0	22	388	Vigorous-intensity activity (≥6 METs)	90.4%	98.7%
			ActiGraph VM (right wrist)					
Actiwatch (right wrist)	Sedentary (< 1.5 METs)	355	76	1	0			
	Light activity (1.5-2.99 METs)	55	368	43	1	Light activity or above (≥1.5 METs)	93.2%	86.2%
	Moderate-intensity activity (3-5.99 METs)	2	114	201	47	Moderate-intensity activity or above (≥3 METs)	92.2%	87.5%
	Vigorous-intensity activity (≥6 METs)	0	6	31	250	Vigorous-intensity activity (≥6 METs)	83.9%	97.0%
			ActiGraph VM (left wrist)					
Actiwatch (left wrist)	Sedentary (< 1.5 METs)	191	119	3	9			
	Light activity (1.5-2.99 METs)	45	322	33	4	Light activity or above (≥1.5 METs)	88.6%	79.6%
	Moderate-intensity activity (3-5.99 METs)	4	155	192	84	Moderate-intensity activity or above (≥3 METs)	90.9%	79.6%
	Vigorous-intensity activity (≥6 METs)	0	14	21	190	Vigorous-intensity activity (≥6 METs)	66.2%	96.8%

AUC: Area under the receiver operating characteristic (ROC) curve; MET: Metabolic equivalent task; VM: vector magnitude.

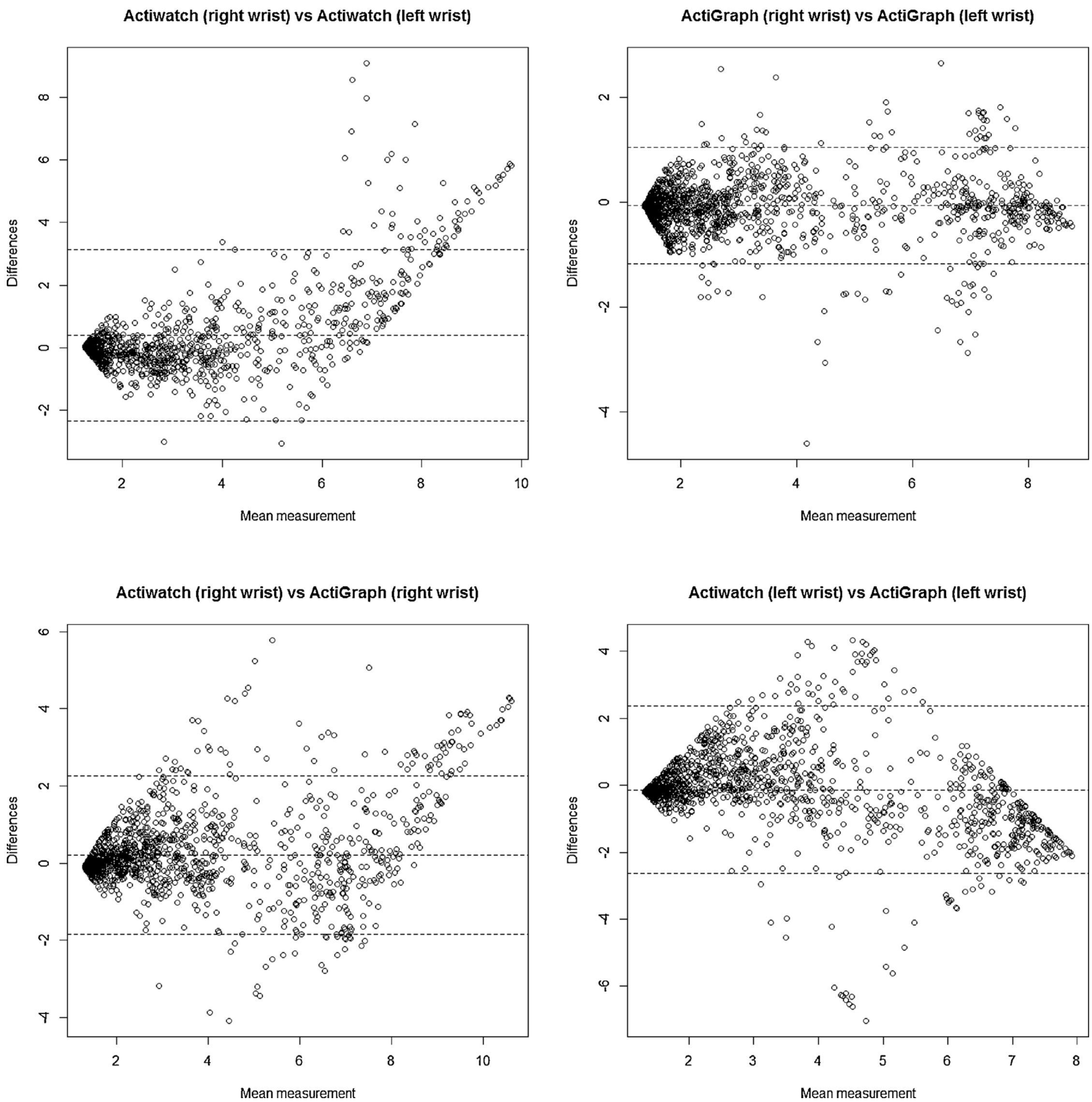

Fig. 3. Bland-Altman plots for Actiwatch 2 and ActiGraph wGT3X-BT.

The plateau phenomenon could be partially attributed to the functional limitations of the accelerometer. A study of some older models of ActiGraph (7164 and GT1M) with limited amplitudes (up to 2.5 *g*) and frequencies (up to 10 Hz) showed no significant difference in cpm for running with a speed ranging from 10 to 20 km/h [20]. Here, we used the latest version of the ActiGraph (wGT3X-BT) with amplitudes up to 8 *g* and frequencies up to 100 Hz, which should have allowed them to overcome the plateau phenomenon [21]. However, our subjects were not trained athletes and we were unable to test the performance of ActiGraph wGT3X-BT with speeds beyond 12 km/h.

Our study was not without limitations. Because our sample is relatively homogeneous, the calibrated equations cannot (and should not) be generalized to other populations, especially older populations that have a different body movement – energy expenditure association from younger populations [22]. In this study, the subjects were required to walk and run on a treadmill, so the swinging motions of the left and right wrist during the experiment were relatively symmetric. This was also reflected in the strong correlations between the left and right accelerometers. Therefore, our results may not be able to generalize to other types of activities with unbalanced wrist movements, for example window cleaning and playing badminton [23]. The accuracy of the calibration equation under free-living condition is to be determined. Similarly, the correlations between wrist-worn Actiwatch 2 and ActiGraph GT3X demonstrated in this study may not (and should not) be compared with the correlations between waist-worn accelerometers, since wrist-worn accelerometers recorded the arm swing movement that provide different cpm patterns with waist movement. Here, we pooled male and female subjects together because another ActiGraph calibration study among college students aged 18–25 showed no sex difference [12]. We have also confirmed this by examining the effect of sex in the calibration equations. However, sex differences were found in an older population (e.g., [22]) and future calibration studies with populations beyond ours are warranted.

5. Conclusion

Our study showed that wrist-worn Actiwatch 2 and ActiGraph wGT3X-BT accelerometers were strongly correlated in PA assessment. Given that the ActiGraph wGT3X-BT has been widely tested as a valid measurement of physical activity, the Actiwatch 2 can also be used for 24-hour activity measurements, as supported by the results of our current study and previous studies on its validity in sleep assessment.

Funding

The Food and Health Bureau of the Hong Kong Special Administrative Region, China provided financial support in the form of Health and Medical Research Fund (Ref 12131741). The sponsor had no role in the design or conduct of this research.

Conflict of interest

None.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.gaitpost.2018.11.023.

References

[1] G. Plasqui, K.R. Westerterp, Physical activity assessment with accelerometers: an evaluation against doubly labeled water, Obesity 15 (2007) 2371–2379.
[2] L.J. Meltzer, C.M. Walsh, J. Traylor, A.M.L. Westin, Direct comparison of two new actigraphs and polysomnography in children and adolescents, Sleep 35 (2012) 159–166.
[3] E. Johannsen, U. Ekelund, H. Hero, C. Marcus, M. Hagströmer, Calibration and cross-validation of a wrist-worn Actigraph in young preschoolers, Pediatr. Obes. 10 (2014) 1–6.
[4] E. Johansson, L.-M. Larisch, C. Marcus, M. Hagströmer, Calibration and validation of a wrist- and hip-worn Actigraph accelerometer in 4-year-old children, PLoS One 11 (2016) e0162436.
[5] M.F. Hjorth, J.-P. Chaput, C.T. Damsgaard, S.-M. Dalskov, K.F. Michaelsen, I. Tetens, et al., Measure of sleep and physical activity by a single accelerometer: can a waist-worn Actigraph adequately measure sleep in children? Sleep Biol. Rhythms 10 (2012) 328–335.
[6] P.H. Lee, D.J. Macfarlane, T.H. Lam, S.M. Stewart, Validity of international physical activity questionnaire short form (IPAQ-SF): a systematic review, Int. J. Behav. Nutr. Phys. Act. 8 (2011) 115.
[7] M.J. Lambiase, K. Pettee Gabriel, L.H. Kuller, K.A. Matthews, Temporal relationships between physical activity and sleep in older women, Med. Sci. Sports Exerc. 45 (2013) 2362–2368.
[8] P.H. Lee, L.K.P. Suen, The convergent validity of Actiwatch 2 and ActiGraph Link accelerometers in measuring total sleeping period, wake after sleep onset, and sleep efficiency in free-living condition, Sleep Breath. 21 (2017) 209–215.
[9] B.E. Ainsworth, W.L. Haskell, S.D. Herrmann, N. Meckes, D.R.J. Bassett, C. Tudor-Locke, et al., Compendium of physical activities: a second update of codes and MET values, Med. Sci. Sports Exerc. 43 (2011) (2011) 1575–1581.
[10] H.C. Pinnington, P. Wong, J. Tay, D. Green, B. Dawson, The level of accuracy and agreement in measures of F_EO_2, F_ECO_2, and $\dot{V}_E$ between the Cosmed K4b^2 portable, respiratory gas analysis system and a metabolic cart, J. Sci. Med. Sport 4 (2001) 324–335.
[11] V.P. Lopes, P. Magalhães, J. Bragada, C. Vasques, Actigraph calibration in obese/overweight and type 2 diabetes mellitus middle-aged to old adult patients, J. Phys. Act. Health 6 (2009) S133–S140.
[12] L.A. Kelly, D.G. McMillan, A. Anderson, M. Fippinger, G. Fillerup, J. Rider, Validity of actigraphs uniaxial and triaxial accelerometers for assessment of physical activity in adults in laboratory conditions, BMC Med. Phys. 13 (2013) 5.
[13] D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4, J. Stat. Soft. 67 (2015) 1–48.
[14] M. Hildebrand, V.T. van Hees, B.H. Hansen, U. Ekelund, Age group comparability of raw accelerometer output from wrist- and hip-worn monitors, Med. Sci. Sports Exerc. 46 (2014) 1816–1824.
[15] J. Staudenmayer, S. He, A. Hickey, J. Sasaki, P. Freedson, Methods to estimate aspects of physical activity and sedentary behavior from high-frequency wrist accelerometer measurements, J. Appl. Physiol. 119 (2015) 396–403.
[16] A.M. McGarty, V. Penpraze, C.A. Melville, Calibration and cross-validation of the ActiGraph wGT3X+ accelerometer for the estimation of physical activity intensity in children with intellectual disabilities, PLoS One 11 (2016) e0164928.
[17] D. John, J.E. Sasaki, J. Staudenmayer, M. Mavilla, P.S. Freedson, Comparison of raw acceleration from the GENEA and ActiGraph™ GT3X+ activity monitors, Sensors 13 (2013) 14754–14763.
[18] J.C. Brønd, L.B. Andersen, D. Arvidsson, Generating ActiGraph counts from raw acceleration recorded by an alternative monitor, Med. Sci. Sports Exerc. 49 (2017) 2351–2360.
[19] J.E. Sasaki, D. John, P.S. Freedson, Validation and comparison of ActiGraph activity monitors, J. Sci. Med. Sport 14 (2011) 411–416.
[20] D. John, B. Tyo, D.R. Bassett, Comparison of four ActiGraph accelerometers during walking and running, Med. Sci. Sports Exerc. 42 (2010) 368–374.
[21] C.V. Bouten, K.T. Koekkoek, M. Verduin, R. Kodde, J.D. Janssen, A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity, IEEE Trans. Biomed. Eng. 44 (1997) 136–147.
[22] N. Zisko, T. Carlsen, Ø Salvesen, N.P. Aspvik, J.E. Ingebrigtsen, U. Wisløff, et al., New relative intensity ambulatory accelerometer thresholds for elderly men and women: the Generation 100 study, BMC Geriat. 15 (2015) 97.
[23] D.W. Esliger, A.V. Rowlands, T.L. Hurst, M. Catt, P. Murray, R.G. Eston, Validation of the GENEA accelerometer, Med. Sci. Sports Exerc. 43 (2011) 1085–1093.