

## Physical activity monitoring devices: energy expenditure comparison in a setting of free-living activities

Received 12<sup>th</sup> August 2019  
Accepted 06<sup>th</sup> October 2019

[www.ijpefs.com](http://www.ijpefs.com)

Matteo Vandoni <sup>a,\*</sup>, Vittoria Carnevale Pellino <sup>a</sup>, Stefano Dell'Anna <sup>a</sup>, Elena Ricagno <sup>a</sup>, Giulia Liberali <sup>a</sup>, Claudia Bonfanti <sup>a</sup>, Luca Correale <sup>a</sup>

<sup>a</sup>Laboratory of Adapted Motor Activity (LAMA), Department of Public Health, Experimental Medicine and Forensic Science, University of Pavia, via Forlanini 2, 27100, Pavia, Italy

\*Corresponding Author Ph: +39 3383440028; Email: [matteo.vandoni@unipv.it](mailto:matteo.vandoni@unipv.it)

**Abstract:** The aim of this study was to evaluate the validity of Energy Expenditure (EE) estimation provided by 3 wearable devices [Fitbit-One (FO), Sensewear Armband (AR) and Actiheart (AC)] in a setting of free-living activities. 43 participants (24 females; 23.4±4.5yrs) performed 9 activities: sedentary (watching video, reading), walking (on treadmill and outdoor), running (on treadmill and outdoor) and moderate-to-vigorous activities (Wii gaming, taking the stairs and playing football). Mean Absolute Percentage Error (MAPE) and Pearson's correlation were calculated to assess the validity of each instrument in comparison to a portable metabolic analyser (PMA). In overall comparison MAPE's were 7,7% for AR (r=.86; p<.0001), 8,6% for FO (r=.69; P<.001), and 11.6% for AC (r=.81; p<.0001). These findings support the accuracy of the wearables. The AR was the most accurate in the whole protocol. However, MAPE results suggest that devices algorithms should be improved for better measure of EE during moderate-to-vigorous activities.

**Key Words:** Energy Expenditure, Physical Activity, Devices, Accuracy, Wearables



**Matteo Vandoni. Ph.D.** in Biomedical and methodological aspects of adapted physical activities at the University of Roma, "Foro Italico". Master's degree in Sport Science at the University of

Pavia. Head of the Laboratory of Adapted Motor Activity (LAMA) of the Department of Public Health, Experimental Medicine and Forensic Science, University of Pavia. His research interests focus on adapted exercise on chronic and metabolic diseases as Diabetes, Obesity and, Multiple Sclerosis.

Public Health, Experimental Medicine and Forensic Science, University of Pavia. Her research interests focus on the biomechanical evaluation of walk and sports performance.



**Stefano Dell'Anna.** Master's degree in Preventive and Adapted Physical Activity at the University of Pavia. He is working as trainer in the Laboratory of Adapted Motor Activity (LAMA) of the Department of Public Health, Experimental Medicine and Forensic Science, University of Pavia. His research interests focus on the sports performance.



**Vittoria Carnevale Pellino Ph.D.** student in Technologies for rehabilitation and sport medicine at the University of Roma, "Tor Vergata". Master's degree in Preventive and Adapted Physical Activity at the University of Pavia. Research Fellow in the Laboratory of Adapted Motor Activity (LAMA) of the Department of



**Elena Ricagno.** Master's degree in Preventive and Adapted Physical Activity at the University of Pavia. She is working as trainer in the Laboratory of Adapted Motor Activity (LAMA) of the Department of Public Health,

*Experimental Medicine and Forensic Science, University of Pavia. She is an expert in adapted exercise for Diabetes and Obesity.*



**Giulia Liberali.** Master's degree in Preventive and Adapted Physical Activity at the University of Pavia. Research Fellow and trainer of the Laboratory of Adapted Motor Activity (LAMA) of the Department of Public Health, Experimental Medicine and Forensic Science, University of Pavia. Her research interests focus on the

assessment of physical activity level and barriers to exercise in people with severe diseases. She is an expert in adapted exercise for Multiple Sclerosis.



**Claudia Bonfanti.** Master's degree in Preventive and Adapted Physical Activity at the University of Pavia. She is working as trainer in the Laboratory of Adapted Motor Activity (LAMA) of the Department of Public Health, Experimental Medicine and Forensic Science, University of Pavia. Her research

interests focus on the sports performance.



**Luca Correale.** Ph.D. in Psychology, Neuroscience and Data Science and a master's degree in Preventive and Adapted Physical Activity at the University of Pavia. Lab Manager and Research Fellow in the Laboratory of Adapted Motor Activity

(LAMA) of the Department of Public Health, Experimental Medicine and Forensic Science, University of Pavia. His research interests focus on exercise effects on diseases as Multiple Sclerosis, Diabetes and, Obesity.

## 1. Introduction

Physical activity (PA) has a fundamental role in human health and help to prevent cardiovascular disease, some cancers, osteoporosis, type-2 diabetes, anxiety and depression[1, 2]. However, adults tend to be less active than the guidelines prescriptions, in fact in Italy only 1 on 3 adults use to practice PA or sport during leisure time and 42% of total population is completely inactive.

Wearable devices for PA are light and affordable and during the last years the availability in commerce increased considerably [3]. Thanks to monitor immediate feedbacks or mobile and internet applications, they give to users various information parameters such as step count, calories burned, time spent in active or sedentary activities and covered distance. Use of objective methods to quantify PA is worldwide increasing as a research and consumer-based tool and it could encourage people to practice more PA[4, 5]. In fact, according to Giroir et al [6] and the PA Guidelines for Americans [7], the use of wearable devices is a new strategy to enhance PA level in people.

In research and field base studies, wearables were tested in free-living condition and in sedentary activities comparing a selection of consumer-level devices and commonly used research-grade accelerometers [8].

However, little is known about the accuracy of these devices during each task at different intensity. Since accuracy measures derive from the mean of the absolute error of the whole protocol, it could be likely that devices output measures have discordant trends during tasks at different intensities, in particular during different type of activities like walking and running both in indoor and in outdoor conditions that could lead interesting data.

Especially in this last condition devices should be preferred to subjective methods such as questionnaires. In fact, it was suggested that these measurement tools underestimate actual sedentary time and the degree of underestimation is extremely variable between subjects [9].

Wearable devices provide also information about the energy expenditure (EE). Usually, in research EE is measured by a gas analyser to avoid the possible lack of accuracy from wearable devices.

In a sort of best practice for the use of wearables, Fredson et al. recommended to calibrate and validate research devices with appropriate protocols and identified the correct use of consumer-based sensors like a challenge for the future of PA studies.

Following these recommendations, Lee, Kim and Welk [10] verified the validity of eight consumer-level devices to assess EE in healthy young adults. They performed a laboratory set of activities lasted 69 minutes, the consumer-level devices were compared against a metabolimeter for oxygen consumption used as the golden standard tool for indirect calorimetry. Researchers ranked devices based on percent of accuracy ranging from 76.5% to 90.7%. To classify instruments, it was considered the whole protocol EE that included sedentary, moderate and vigorous intensity activities[10].

Thus, the aim of this study was to evaluate the validity of EE estimation provided by 3 wearable devices in a setting of both laboratory and free-living activities in comparison to a portable metabolic analyser (PMA) during tasks at different intensity.

## 2. Materials and Methods

### 2.1 Participants

Forty-three subjects (24 females; 23.4±4.5 years; mean±sd) voluntarily took part to the study. Participants did not have diseases or illnesses and did not use drugs that would affect their body weight or metabolism. Subjects were recruited from University of Pavia students through internet announcements.

Approval from the academic review board of Kinesiology Course was obtained before beginning this study. Participants were aware of the procedures and purpose of the study before they signed the informed consent document.

### 2.2 Measures

Many wearable devices were produced and used in order to provide an objective indicator of PA such as the EE, which represents a fundamental factor to quantify PA. This study aimed to examine the validity of EE, analysed through devices in different types of activities (sedentary, walking, running and moderate to vigorous), that could represent a setting of real-life PA.

Moreover, it was examined more deeply the accuracy of accelerometers in specific conditions showing data that could validate the values of EE output with a PMA (K4b2 COSMED, Rome, Italy), which represented the gold standard of the study.

Before testing, anthropometric data (height, weight and BMI), the age of participants and basal metabolism of each subject were measured. (table 2); at last, all devices were prepared with the personalized data of each subject to make the software of each instruments more accuracy.

### 2.3 Procedures

The protocol of the study was composed by 9 different activities and lasted totally 64 min (fig. 1). The order of activities intensity is incremental to avoid effects of fatigue and to facilitate the succession of indoor to outdoor trials: all the subjects began with indoor trials as reading a newspaper and watching a video, then they performed both walking and running on treadmill and afterward the subjects went up and down on stairs and played with Nintendo Wii gaming. At least, all the subjects continued with outdoor activities like walking, running and football small match. While the intensity of work on treadmill were the same for all the subjects, walking and running outdoor were at self-selected gait. Every activity was performed for 5 minutes. There was a 1-minute rest between each activity to facilitate transitions and the tracking of data. All the activities were classified into four distinct PA intensities: 1) Sedentary (reading a newspaper, watching a video) 2) Walking (treadmill walking at 4km/h, self-paced over ground walking), 3) Running (treadmill running at 10km/h, self-paced over ground running) and 4) Moderate-to-vigorous

activities (going up and down the stairs, Wii dance play and playing football with the researchers).



**Figure 1.** The order of different activities performed by all the subjects in the study. Different colours define the four different tasks intensities

**Table 1** Table 1 Description of protocol devices: their allocation on the body (Positioning), devices typology (Sensor), their measurements (Parameter), the battery length (Range) and data transmission modality (Interface).

Instruments	K4b2 (COSMED, Italy)	Fitbit One (Fitbit Inc., USA)	Actiheart (CamNtech Inc, England)	Sensewear Armband (BodyMedia, USA)
Figure				
Positioning	Chest-face	Hip	Chest	Triceps
Sensor	Analyzer O2,CO2, Barometer, GPS	Accelerometer 3 axes	Accelerometer 3 axes, HR monitor	Thermometer, Accelerometer 2 axes
Parameter	VO2, VCO2	Step, Kcal, EE, Dist, H&Q of sleep	HR, EE	EE, Step, Skin, Temp.
Range	6 hours	10-14 days	21 days	12 days
Interface	Display, Holter, Telemetry	Display LCD, Website	Software	Software

HR= Heart Monitor; EE=Energy Expenditure; Dist=Distance; H&Q sleep= Hours and Quality of sleep; Temp= Temperature

Most of the wearables did not provide direct access to the raw data; therefore, valuations of EE were obtained directly from the associated software for each device.

The participants were fitted with the PMA and three different types of wearables. PMA was worn on the chest and the back, whereas SenseWear Armband (AR) (Bodymedia, Pittsburgh, USA) was worn on the non-dominant arm. Fitbit One (FO) (Fitbit Inc., San Francisco, CA) was worn on the hip and and Actiheart (AC) (CamNtech Inc, England) was worn on the chest. All these instruments were synchronized and initialized using the participant's personal information (age, gender, height, weight, handedness and smoker/non-smoker) before every measurement. The characteristics of each instruments is described in table 1.

### 2.4 Analysis

Data are shown as mean  $\pm$  standard deviation, minimum and maximum values were recorded as range. Breath by breath values from PMA were aggregated to provide minute by minute mean to simplify comparison with accelerometers data. EE of both wearables devices and PMA was calculate during the entire monitoring period and each single task. Resting period between the activities were not evaluated. Statistical analysis aim was to compare EE of every wearable devices with PMA values (criterion measure). Pearson's correlations were calculated to analyse both overall group level and single tasks associations.

Finally, mean absolute percent error (MAPE), calculated as the average of absolute differences between wearable devices and PMA value divided by PMA value and multiplied by 100.

### 3. Results

The anthropometric characteristics of the participants are reported in table 1. Men aged between 19 and 38 years old. Medium BMI was  $23,1 \pm 1,9$  for men and  $21,4 \pm 2,7$  for women.

Figure 2 to figure 6 represent the MAPE of every device relative to the PMA. Protocol's results were divided in four different categories based on different intensities: sedentary activities (Figure 4), walking activities (Figure 5), running activities (Figure 6) and moderate to vigorous activities (Figure 7). Figure 2 shows MAPE of the whole protocol: AR shows the lowest error (7,7%), followed by FO (8,6%), AC (11,6%). Every wearable device shows similar MAPE in sedentary activities between 12,0% and 13,0%. Conversely AR reveals 13,3% MAPE in walking activities, followed by AC (19,7%) and FO (50,0%).

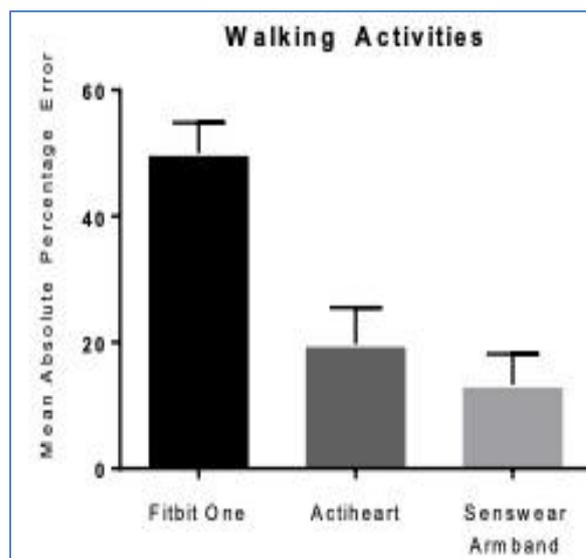
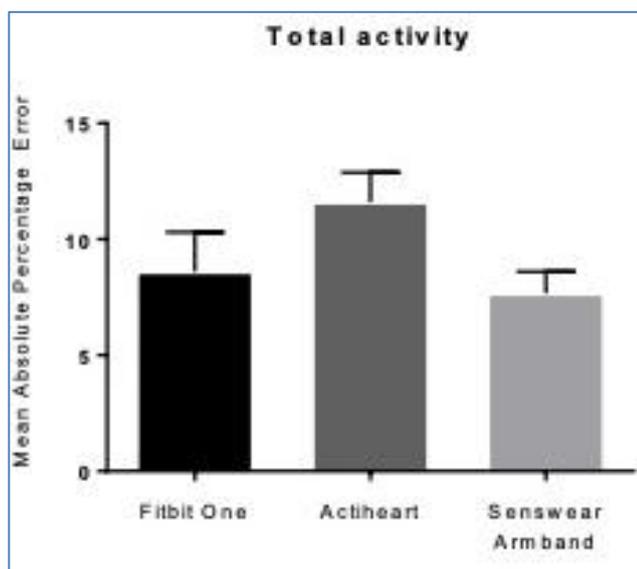
This discrepancy is reduced in running activities, where MAPE's range between 13,0% (AR) and 18,4% (AC). FO is the most accurate device in moderate to vigorous activities showing 16,4% of error, AR's MAPE is 26,4% and AC's MAPE is 41,2%.

**Table 2** Anthropometric characteristics of the participants in the protocol

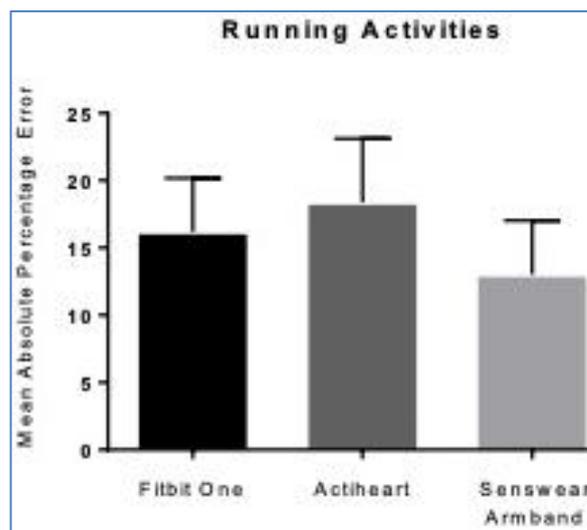
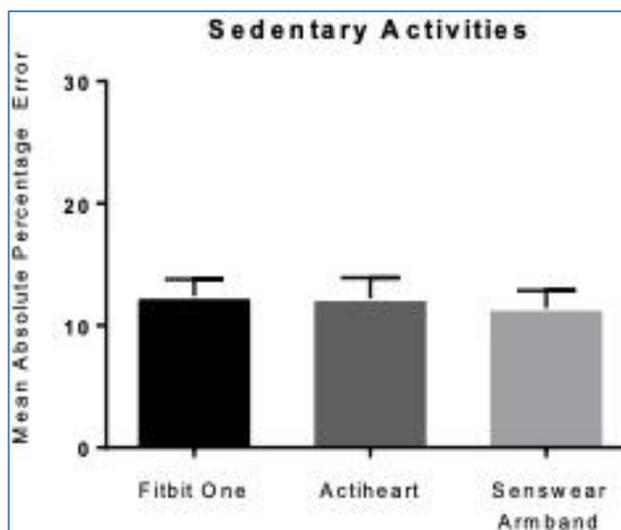
	MEN		WOMEN	
	Mean $\pm$ SD	Range	Mean $\pm$ SD	Range
AGE	23,4 $\pm$ 4,5	19,0 – 38,0	22,0 $\pm$ 3,2	19,0 – 30,0
HEIGHT (cm)	178,3 $\pm$ 6,9	170,0 – 190,0	168,4 $\pm$ 5,5	158,0 – 185,0
WEIGHT (kg)	72,4 $\pm$ 9,0	56,0 – 86,0	59,9 $\pm$ 11,6	49,0 – 80,0
BMI	23,1 $\pm$ 1,9	19,4 – 26,3	21,4 $\pm$ 2,7	17,8 – 29,4

**Table 3** Table 3 – Mean ( $\pm$ SD) and Range of estimated energy expenditure (kcal) of each wearable device. PMA measured  $357 \pm 65,17$  kcal (mean  $\pm$  SD) and devices ranged between 337 (AC) and 492 kcal (AR). AC’s sample is significantly reduced because of missing data during high intense activities.

	N	Mean $\pm$ SD	Range
PMA	39	$357,00 \pm 65,17$	221 - 486
FO	39	$350,13 \pm 49,68$	267 - 484
AC	27	$337,11 \pm 77,12$	250 - 466
AR	37	$354,97 \pm 59,27$	250 - 492

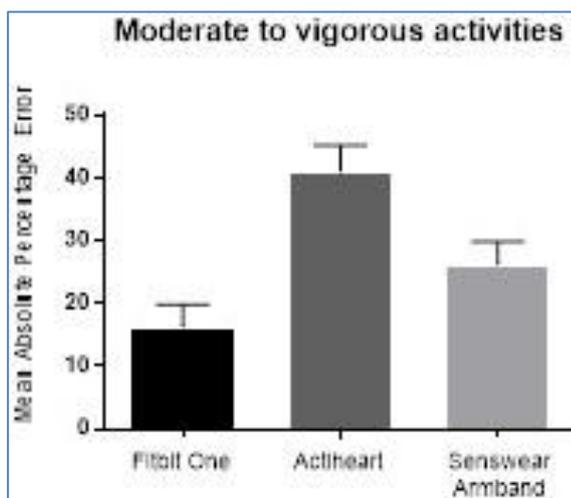


**Figure 2.** MAPE ( $\pm$ SEM) for all devices in the whole protocol. **Figure 4.** MAPE ( $\pm$ SEM) in walking activities.



**Figure 3.** MAPE ( $\pm$ SEM) in sedentary activities.

**Figure 5.** MAPE ( $\pm$ SEM) in running activities.



**Figure 6.** MAPE (±SEM) in moderate to vigorous activities.

Table 4 and table 5 show correlation coefficient (r) between standard values (PMA) and the others wearable devices in different intensities of exercise. Correlation results are divided in the same categories as MAPEs. AR seems to be the device with the strongest correlation with PMA in every kind of activity and shows important correlation values with AC. Other values range from 0,448 to 0,865.

**Table 4.** Correlation coefficient in total activity protocol

	PMA	FO	AC	AR
PMA	1	0,695**	0,813**	0,861**
FO		1	0,760**	0,747**
AC			1	0,848**
AR				1

*\*\*Correlation is significant at 0,05 (two tailed).*

**Table 5** Correlation coefficient in all the activities performed in the protocol. *\*\*Correlation is significant at 0,05 (two tailed).*

PMA	FO	AC	AR	PMA	Fo	Ac	Ar	
1	0,629**	0,740**	0,816**	PMA	1	0,448**	0,547**	0,722**
0,642**	1	0,715**	0,789**	FO	0,534**	1	0,553**	0,571**
0,628**	0,563**	1	0,781**	AC	0,619**	0,679**	1	0,740**
0,718**	0,503**	0,865**	1	AR	0,717**	0,585**	0,648**	1

Legend of different activities:

- Sedentary activities
- Walking activities
- Moderate to vigorous activities
- Running activities

## 4. Discussion

The aim of this study was to evaluate the validity of EE estimation in healthy adults provided by 3 wearable PA devices in a setting of both laboratory and free-living activities in comparison to PMA during tasks at different intensities. Despite the accuracy of the EE measures of various wearables has already been analysed [10, 11], there's poor information about the precision of these devices during activities at different intensities (from sedentary to high intensity).

All wearable devices showed appropriate results data for the whole protocol evaluation (MAPE range approximatively from 8 to 12%). The error rates in the present study were comparable (< 12%) to results presents in other study[10, 12], this means that the wearable devices are providing similar accuracy as the PMA. In particular, we found that AR provided the better correlations with PMA in the whole protocol ( $r = 0.86$ )[10] and the best MAPE (7.7%). FO and AC showed good results, respectively 8.6% and 11.6% for MAPE values and correlation of 0,69 and 0,81 in the overall protocol.

Every device shows similar MAPE in sedentary activity between 12,0% and 13,0%. This result could be mainly attributed to the algorithm for EE in basal condition. In walking activities AR had the best performance with 13,3% of MAPE, followed by AC (19,7%) and FO (50,0%). We believed that FO have worse MAPE values because of its placement on the belt, this position could be not optimal to evaluate the walk.

During the walk we observed the major discrepancy of values about all devices. The best correlation coefficient in walking activities is found in AR with  $r = 0,72$ . This difference decreased in running activities, where MAPE's range between 13,0% (AR) and 18,4% (AC).

Concerning moderate to vigorous activities, FO was the most accurate device showing 16,4% of error, AR is at 26,4%. AC showed high MAPE's value because it didn't record continuously during vigorous activities.

In light of these results, AR is probably the best device to measure activity at different tasks,

while to measure only vigorous activity is better to use the FO.

The strength of this study was to compare the EE from the devices during different intensity task, the protocol was designed to include typical activities that would be reflective of normal adult behaviour. Previous research [13] showed higher correlations with O2 ( $r = 0.85 - 0.93$ ) under laboratory conditions in contrast to lower correlations in free-living conditions ( $r = 0,48 - 0,59$ ). Daily activities include a considerable amount of upper-body movements that not may be captured by the accelerometer-based devices, so it's possible that the device overestimated some activities and underestimated others [3, 10]. However, the EE estimation provided by the 3 wearable PA devices used in this study was reasonable. In fact, all the devices had high correlation values in total activity protocol (FO  $r = 0,70$ ; AC  $r = 0,81$ ; AR  $r = 0,86$ ) in accordance with results of Lee at al [10].

Probably, the use of PA wearable devices placed on the arm avoid this problem during free-living activities, in fact the AR was the most accurate in this study.

## 5. Conclusion

In conclusion, the present study supports the accuracy of the three devices used to estimate EE in healthy adults, especially considering the different intensities during the protocol. In particular AR result to be the most accurate in the whole protocol, including walking and running activity. However, the findings regarding MAPE suggest that the internal devices algorithms should be improved for better measure of EE during different tasks, in particular in moderate to vigorous activities. Considering that some bias already exist in EE estimation, these results add new knowledge for specific activities evolution and will help researchers to better use the right device for the peculiar setting of the study.

## References

- [1] S.N. Blair and J.C. Connelly, How much physical activity should we do? The case for moderate amounts and intensities of physical activity,

- Research Quarterly for Exercise and Sport*, 67 (1996) 193–205.
- [2] S. J. Petruzzello, D.M. Landers, B.D. Hatfield, K. A. Kubitz, and W. Salazar, A meta-analysis on the anxiety-reducing effects of acute and chronic exercise. Outcomes and mechanisms, *Sports Medicine*, 11 (1991) 143–182.
- [3] S.R. Gomersall, N. Ng, N.W. Burton, T.G. Pavey, N.D. Gilson, and W.J. Brown, Estimating Physical Activity and Sedentary Behavior in a Free-Living Context: A Pragmatic Comparison of Consumer-Based Activity Trackers and ActiGraph Accelerometry, *Journal of Medical Internet Research*, 18 (2016) e239.
- [4] K.R. Evenson, M.M. Goto, and R.D. Furberg, Systematic review of the validity and reliability of consumer-wearable activity trackers, *International Journal of Behavioral Nutrition and Physical Activity*, 12 (2015) 159.
- [5] M. Vandoni, L. Correale, M. Del Bianco, L. Marin, and E. Codrons, Does reactivity to accelerometers occur in a single trial? Brief report in a sample of young adults, *Journal of Health Psychology*, 22 (2017) 1458–1462.
- [6] B. P. Giroir and D. Wright, Physical Activity Guidelines for Health and Prosperity in the United States, *JAMA*, 320 (2018) 1971.
- [7] K. L. Piercy *et al.*, The Physical Activity Guidelines for Americans, *JAMA*, 320 (2018) 2020.
- [8] T. Ferguson, A. V Rowlands, T. Olds, and C. Maher, The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study, *International Journal of Behavioral Nutrition and Physical Activity*, 12 (2015) 42.
- [9] B. K. Clark, T. Sugiyama, G. N. Healy, J. Salmon, D. W. Dunstan, and N. Owen, Validity and reliability of measures of television viewing time and other non-occupational sedentary behaviour of adults: a review, *Obesity Reviews*, 10 (2009) 7–16.
- [10] J.-M. Lee, Y. Kim, and G.J. Welk, Validity of Consumer-Based Physical Activity Monitors, *Medicine & Science in Sports & Exercise*, 46 (2014) 1840–1848.
- [11] R.K. Reddy, R. Pooni, D.P. Zaharieva, B. Senf, J. El Youssef, E. Dassau, F.J. Doyle Iii, M.A. Clements, M.R. Rickels, S.R. Patton, J.R. Castle, M.C. Riddell, P.G. Jacobs, Accuracy of Wrist-Worn Activity Monitors During Common Daily Physical Activities and Types of Structured Exercise: Evaluation Study, *JMIR mHealth uHealth*, 6 (2018) e10338.
- [12] M.B. Nelson, L.A. Kaminsky, D.C. Dickin, and A.H.K. Montoye, Validity of Consumer-Based Physical Activity Monitors for Specific Activity Types, *Medicine & Science in Sports & Exercise*, 48 (2016) 1619–1628.
- [13] M. McCarthy and M. Grey, Motion Sensor Use for Physical Activity Data: Methodological Considerations, *Nursing Research*, 64 (2015) 320–327.

## Funding

This study was not funded by any grant

## Acknowledgements

The authors would like to thank every participant for his effort and time.

## Conflict of interest

None of the authors have any conflicts of interest to declare.

## Informed consent

All participants gave written informed consent to participate in this study.

## About The License



The text of this article is licensed under a Creative Commons Attribution 4.0 International License