

Article

## A Comparison of Two Motion Sensors for the Assessment of Free-Living Physical Activity of Adolescents

Roman Cuberek <sup>1\*</sup>, Walid El Ansari <sup>2</sup>, Karel Frömel <sup>1</sup>, Krzysztof Skalik <sup>3</sup> and Erik Sigmund <sup>1</sup>

<sup>1</sup> Center for Kinanthropology Research, Faculty of Physical Culture, Palacky University in Olomouc, Czech Republic; E-Mails: karel.fromel@upol.cz (K.F.); erik.sigmund@upol.cz (E.S.)

<sup>2</sup> Faculty of Sport, Health and Social Care, University of Gloucestershire, Gloucester GL2 9HW, UK; E-Mail: walidansari@glos.ac.uk

<sup>3</sup> Academy of Physical Education, Katowice, Poland; E-Mail: k.skalik@awf.katowice.pl

\* Author to whom correspondence should be addressed; E-Mail: roman.cuberek@upol.cz; Tel.: +420 585 636 171; Fax: +420 585 636 104.

*Received: 5 February 2010; in revised form: 16 March 2010 / Accepted: 25 March 2010 /*

*Published: 6 April 2010*

---

**Abstract:** This study assessed and compared the daily step counts recorded by two different motion sensors in order to estimate the free-living physical activity of 135 adolescent girls. Each girl concurrently wore a Yamax pedometer and an ActiGraph accelerometer (criterion measure) every day for seven consecutive days. The convergent validity of the pedometer can be considered intermediate when used to measure the step counts in free-living physical activity; but should be considered with caution when used to classify participants' step counts into corresponding physical activity categories because of a likelihood of 'erroneous' classification in comparison with the accelerometer.

**Keywords:** pedometer; accelerometer; validity; monitoring; measurement; step counter; categorization of physical activity

---

## **1. Introduction**

In order to assess the level of physical activity (PA) and evaluate PA interventions [1], measures are required that have high validity and reliability, whilst being practical and nonreactive [2]. Validity is considered the most important attribute, but for PA assessment tools, and particularly for free-living PA, verifying validity is difficult task [3].

PA is the sum of all those partial bodily movements realized by a muscle-skeletal system and resulting in energy expenditure [4]. Using this definition, appropriate measures are required to assess the levels of attainment of such bodily movements. For this purpose, the techniques used in a practice comprise three groups [2]: criterion standards (direct observation, doubly-labeled water, and indirect calorimetry); objective techniques (heart rate, motion sensors: pedometers and accelerometers); and subjective techniques (self-reported questionnaires, interviewer-administered questionnaires, proxy-reporting questionnaires and diaries).

Although the optimal technique of accurate assessment of free-living PA of children and adults is difficult to ascertain, direct observation is still the most acceptable [2]. However, the use of direct observation to assess free-living PA is hindered by that the technique entails considerable time and personnel requirements [5-7]. Hence, questionnaires and/or motion sensors (MS, e.g., accelerometers and pedometers) are frequently employed instead. Such measurement tools have had their validity [2,8-11] and reliability [2,12,13] verified under various conditions. For instance, De Vries and colleagues [14] provided summaries of methodological studies published between 2004 and 2007 that focused on the general characteristics, reproducibility, validity and feasibility of MS used in the assessment of PA (appraisals of three pedometers and nine accelerometers). Based on their analysis, they suggested MS to be sufficiently valid to assess PA in youth, and that most MS that they assessed were reliable enough. They also noted that using MS as an estimator of energy expenditure needed caution, due to the significant under- and overestimation when compared with the doubly-labeled water technique, a method that is considered a “gold standard” of energy expenditure estimation of free-living PA [10]. Although doubly-labeled water is accurate, its drawbacks are that it is expensive; it does not provide information about the intensity, frequency, duration, or type of PA; and it requires specific expertise [15]. Consequently, the last decade has witnessed an increase in the use of MS (pedometers, accelerometers) as PA assessment tools [14].

The pedometer is a simple MS device based on bodily acceleration or deceleration in one direction. It enables the observer to measure the number of steps taken over a period of time. The disadvantages of this technique are the inability to measure the intensity of PA; record the amount of PA (in terms of step counts) over several time periods; and document any energy expenditure increase achieved through carrying a load or walking/running uphill [3,16,17]. Nevertheless, researchers [14,17] have signified pedometers to be sufficiently valid and reliable to assess free-living PA in cases where the intensity of PA is not required.

The accelerometer is another type of MS used to assess PA; it is more sophisticated and provides more detailed outputs in comparison with pedometers. However, when compared with pedometers, the disadvantage of accelerometers is their price, which can be a limiting factor particularly for wide-scale investigations that recruit many participants. Hence pedometers have become an appropriate alternative when researchers seek to measure only the number of steps achieved [5].

As regards to MS, the Yamax Digi-walker is the most verified pedometer while the ActiGraph is the most verified accelerometer [14]. MS have the additional advantages of being light in weight and relatively low priced when compared with objective techniques (indirect calorimetry or doubly-labeled water) [18]; and that concurrently within their development they were validated and calibrated in various conditions.

As the internal mechanisms of such MS devices are different, the reliability and validity of each type of device is also different. Although such differences in qualities were of less significance in controlled conditions [13], more often such differences in qualities were important when free-living PA was being assessed [6,11,19-21]. Thus it is important to ascertain the extent of agreement (and disagreement) between pedometers and accelerometers in documenting the levels of free-living PA in different settings. In order to address this gap, the present study employed a sample of free-living adolescents in Poland and compared the Yamax Digi-walker (the most verified pedometer) and the ActiGraph (the most verified accelerometer).

### *1.1. Aims of the Study*

This study assessed and compared, in a sample of free-living adolescents (N = 135), the daily step counts recorded by two different MS that operate with different principles and are both commonly employed in measuring the number of steps. We appraised the level of agreement of step counting of two MS: pedometer (Yamax Digi-walker SW-701) and accelerometer (criterion measure; ActiGraph GT1M). The specific objectives were to:

- assess the step counts recorded by each of the two MS for the same participant everyday over a period of one week;
- compare the step counts recorded by each of the two MS for the same participant in order to detect any inter-device differences on each day over a period of 1 week, employing the accelerometer as the criterion measure; and,
- classify the step counts recorded by each of the two MS for the same participant into PA categories, and compare these PA categories for the same participant in order to detect any inter-device differences in the categories on each day over a period of 1 week.

## **2. Methods**

### *2.1. Participants, Settings and Procedures*

As only 149 pairs of MS (pedometers and accelerometers) were available for the study, and we wanted to measure all participants in one 'batch' at one time, this limited the maximum sample size of participants that could be recruited to 149 pupils (mean age  $18.0 \pm 0.6$  years; BMI range:  $16.5\text{--}29.3 \text{ kg}\cdot\text{m}^{-2}$ ). Hence only four high schools were selected to implement the study, and the selection was based on successful previous collaborations between the Center of Kinanthropology Research at Palacky University (Czech Republic) and the schools. The schools had similar study curricula, and were all in urban areas in the Katowice region in Poland that had similar conditions for PE and free-living PA (Dabrowa, Katowice, Sosnowiec, and Swietochlowice; 64% of participants lived in cities populated by >100,000 inhabitants, 7% lived in cities populated by <30,000 inhabitants,

and the rest of the sample lived in cities that were in between these two extremes). For instance, the four schools provided the same quantity of physical education (three hours of physical education per week), and had similar material facilities for PA (e.g., indoor and outdoor sport and exercise facilities *etc.*). None of the participants was obese or a professional athlete, and none of them had any mobility-impairing disabilities. About 8% of the participants were engaged in organized PA more than four times a week.

The Ethics committee at the participating Universities (in the Czech Republic and in Poland) agreed to the proposed research and approved the study. At the four sites, the aims and objectives of the study were clarified to all pupils who were then invited to participate in the research if they and their parents agreed (signed written informed consent). Participation was voluntary and pupils could withdraw from the study at any time. In order to be enrolled in the study, pupils had to agree to wear two MS devices continuously for the period of monitoring (one week). Equal numbers of participants were randomly selected from classes at each of the four schools. Data were anonymous and confidential, and data protection was observed at all times.

We only selected female pupils from the four schools in agreement with others that girls were more responsive than boys in wearing a pedometer [22]. Hence our sample comprised 149 girls who concurrently wore two MS: a Yamax Digi-Walker SW-701 (Yamax Co., Yasama Corp., Tokyo, Japan) pedometer (PDW), and an ActiGraph GT1M (Manufacturing Technology Inc., FL, USA) accelerometer (AAG) over a 7-day period in October 2008. The accelerometer was set according to the manufacturer's specifications: it was calibrated and programmed to record information in 1-minute epochs (*i.e.*, the normal factory default sensitivity as set by the manufacturer). Participants wore both devices on opposite sides of the waist (located in pouches firmly fixed on elastic fabric belt) for a mean of  $14.06 \pm 2.13$  hours per day (range = 8.98 to 19.77 hours per day), taking it off when sleeping and for hygiene (e.g., bathing, *etc.*). We excluded data from 10 girls due to technical causes (e.g., when any of the two devices had been left lying for some time or when a device was lost); and from another four girls due to record failures (data were considered incorrect if it was out of range of 1,000–30,000 steps per day). Thus, the present analysis employed data from 135 girls only.

## 2.2. Variables (Step Counts per Day; and Categories of PA)

When participants removed their PDW every evening, they recorded the data from the PDW monitor on a special data record sheet that we provided and then reset the pedometer to zero. Data from the AAG was retrieved by the special ActiPA2006 software [23]. Then, four sets of information were retrieved from the original data: (1) daily step counts achieved in each of the seven days of the monitoring week (Monday to Sunday); (2) daily step counts achieved within the whole week; (3) daily step counts achieved within the period of five work-days; and, (4) daily step counts achieved within the period of the weekend. When one of the paired values (readings from one MS) was missing for a given day for a given participant, the participant was excluded from the particular analysis. Then, for each of the two MS we employed, based on the recorded daily step counts, participants were grouped into five categories of PA, in agreement with published threshold guidelines [1]. The categories comprised: 'sedentary' (<5,000 steps/day); 'low active' (5,000–7,499 steps/day); 'somewhat active' (7,500–9,999 steps/day); 'active' (10,000–12,499 steps/day); and, 'highly active' ( $\geq 12,500$  steps/day).

### 2.3. Statistical Analysis

Data was analyzed using STATISTICA 8 (StatSoft Inc., Tulsa, OK, USA) with significance level set at  $p < 0.05$ . As the data was not normally distributed for most variables, descriptive statistics are presented as medians (Mdn) and inter-quartile ranges (IQR). Correlation analyses (Spearman's rank correlation coefficient) quantified the linear relationship between step counts recorded by the pedometer (steps per day<sub>PDW</sub>) and accelerometer (steps per day<sub>AAG</sub>) in all the individual days, in the 2-day period (weekend), in the 5-day period (workdays), and in overall 7-day period (week days). Spearman's rank correlation coefficient was also used to assess the relationship between inter-device differences in step counts and the number of hours participants wore the devices. Kruskal-Wallis ANOVA tested whether inter-device differences were associated with particular schools (undertaken separately for all individual days, 2-day period, 5-day period, and 7-day period).

For the actual number of steps achieved, Wilcoxon-matched-pair test was used to verify non-zero differences between step counting of the two MS. Effect size [24] was then computed in order to assess the 'practical' significance of inter-device differences, employing the formula:  $d = 2 \cdot \frac{Z}{\sqrt{N}}$ , where  $Z$  was calculated from Wilcoxon-matched-paired test that was used to compute statistical significance. When we further classified the actual number of steps achieved into corresponding PA categories, intra-class correlation (ICC) appraised the degree of agreement of participants' PA categorization that were derived from amount of steps obtained from PDW and AAG.

### 3. Results

Kruskal-Wallis ANOVA indicated that inter-device differences in steps counts were independent of the school where participants were recruited from (in all cases  $p$ -levels were at least  $\geq 0.15$ ). Likewise, the correlations between inter-device differences and duration of wearing the devices throughout the day were consistently low, and statistically non-significant in all the individual days, and in the 2-day, 5-day, and 7-day periods ( $r_{sp}$  ranged between 0.07<sub>5-day period</sub> to 0.14<sub>2-day period</sub>; in all,  $p \geq 0.10$ ).

Table 1 depicts that collectively in the total 7-day period of monitoring, depending on the MS, girls accumulated Mdn<sub>AAG</sub> = 8,874 (IRQ = 6,695–10,777) steps per day, and Mdn<sub>PDW</sub> = 8,489 (IRQ = 6,647–11,120) steps per day. For both devices, inter-day differences (steps per day) for individual days were significant in many instances ( $p < 0.05$  in 9 out of the 15 possibilities that existed to compare the individual 7 days). Furthermore, for both MS, the differences between number of steps achieved in the periods 'workdays' and 'weekend' (inter-period differences) were also significant. Inter-device differences (steps per day) were significant only on Wednesday ( $p = 0.025$ ) and Sunday ( $p = 0.036$ ), and there were no inter-device differences for the periods 'workdays', 'weekend', and 'whole week'. For the individual days where the differences between the two MS were significant, the effect size (*i.e.*, 'size' of the difference between the two devices) ranged from 0.36 (Sunday) to 0.39 (Wednesday) (*i.e.*, mostly poor to small effect) (Table 1). Furthermore, none of the differences (in steps per day) between both devices exceeded 1,000 steps per day—a level that might be insufficient for some authors to consider such differences to be of 'practical' significance (we based

‘practical’ significance on the rationale that others [1] employed levels of 2,500 steps per day as differences between successive incremental categories of PA levels).

**Table 1.** Median number of steps achieved: Inter-device differences between pedometer and accelerometer (N = 135).

Period of measurement	Pedometer	Accelerometer	Inter-device difference		
	PDW	AAG	Diff <sub>(median)</sub>	d	p-value
	Mdn (IQR)	Mdn (IQR)			
<b>Individual days</b>					
Monday	8,593 (6,155; 11,546)	8,915 (6,289; 11,455)	322	0.14	0.131
Tuesday	8,372 (6,016; 11,581)	8,565 (5,809; 10,782)	193	0.07	0.684
Wednesday	9,414 (7,212; 11,884)	9,051 (6,876; 11,689)	−363	0.39	0.025*
Thursday	7,176 (4,978; 10,204)	7,527 (5,691; 10,064)	351	0.13	0.434
Friday	9,074 (6,883; 13,581)	10,043 (7,296; 12,622)	969	0.07	0.702
Saturday	8,049 (5,185; 11,804)	8,929 (5,844; 12,186)	880	0.02	0.926
Sunday	6,753 (4,247; 10,610)	6,410 (4,082; 9,722)	−343	0.36	0.036*
<b>Time periods</b>					
5-day period (Workdays)	9,094 (6,910; 10,963)	9,098 (6,927; 11,132)	4	0.15	0.397
2-day period (Weekend)	7,667 (5,372; 10,620)	7,772 (5,607; 10,745)	105	0.18	0.300
7-day period (Whole week)	8,489 (6,647; 11,120)	8,874 (6,695; 10,777)	385	0.14	0.423

PDW: pedometer (Yamax Digi-walker SW-701); AAG: accelerometer (ActiGraph GT1M); Mdn: median; IQR: inter-quartile range; Diff<sub>(median)</sub> = Mdn<sub>AAG</sub> - Mdn<sub>PDW</sub>: inter-device difference of medians; \* statistically significant inter-device difference (Wilcoxon-matched-pair test); d: effect size.

Table 2 shows that across each of the seven individual days, in the 5-day period (workdays), in the 2-day period (weekend), and in the overall 7-days period (whole week), the correlations between the two MS were positive, moderately or slightly high correlated ( $r_{sp}$  range = 0.65 to 0.74), and significant ( $p < 0.001$ ).

We then classified the step counts recorded by each of the two MS for the same participant into categories, and compared these categories for the same participant in order to detect any inter-device differences in the categories.

**Table 2.** Number of steps achieved: Inter-device correlation between pedometer and accelerometer.

Number of steps achieved on individual days	$r_{sp}$	p-value
Monday	0.68	< 0.001
Tuesday	0.66	< 0.001
Wednesday	0.65	< 0.001
Thursday	0.69	< 0.001
Friday	0.72	< 0.001
Saturday	0.69	< 0.001
Sunday	0.74	< 0.001
Time periods		
5-day period (workdays)	0.64	< 0.001
2-day period (weekend)	0.66	< 0.001
7-day period (whole week)	0.65	< 0.001

N = 135;  $r_{sp}$  – Spearman’s rank order correlation coefficient.

Using the pedometer-determined step counts to generate the PA categories, 7.4% of our girls were classified as ‘sedentary’; 25.9% as ‘low active’; 33.3% as ‘somewhat active’; 22.2% as ‘active’; and 11.1% as ‘highly active’ in the whole 7-day period of monitoring. Table 3 illustrates that across each of the seven individual days, the ICC for the participants’ PA categorization based on PDW- and AAG-counted steps, varied from 0.69 to 0.62; and within the 2-day, 5-day, and 7-day periods, the ICC ranged from 0.64 to 0.63.

**Table 3.** Category of PA achieved: intraclass correlation between participants’ PA categorizes based on pedometer and accelerometer step counting (N = 135).

Category of PA achieved on	ICC	95% CI
Individual days		
Monday	0.65	0.54; 0.73
Tuesday	0.63	0.52; 0.72
Wednesday	0.62	0.50; 0.71
Thursday	0.69	0.59; 0.77
Friday	0.69	0.59; 0.77
Saturday	0.68	0.58; 0.76
Sunday	0.69	0.59; 0.77
Time periods		
5-day period (workdays)	0.63	0.51; 0.72
2-day period (weekend)	0.62	0.50; 0.71
7-day period (whole week)	0.64	0.53; 0.73

ICC: intraclass correlation; CI: confidence interval.

**Table 4.** Agreement and disagreement between participants' PA categorization based on pedometer (PDW) and accelerometer (AAG) step counts (N = 135).

Differences in Step Count Categories*	Duration of Monitoring									
	Time Periods			Individual Days						
				Workdays					Weekends	
	7 days	5 days	2 days	Mon	Tue	Wed	Thu	Fri	Sat	Sun
<b>AGREEMENT (PDW = AAG)</b>										
<b>A. TOTAL AGREEMENT</b>	<b>53</b>	<b>43</b>	<b>52</b>	<b>48</b>	<b>52</b>	<b>50</b>	<b>60</b>	<b>54</b>	<b>53</b>	<b>57</b>
<b>DISAGREEMENT (PDW ≠ AAG)</b>										
<b>B. TOTAL DISAGREEMENT</b>	<b>47</b>	<b>57</b>	<b>48</b>	<b>52</b>	<b>48</b>	<b>50</b>	<b>40</b>	<b>46</b>	<b>47</b>	<b>43</b>
<b>Underestimation (PDW &lt; AAG)</b>										
<b>Total % of participants underestimated by PDW†</b>	23	30	21	20	26	20	21	28	21	15
1 category difference	19	24	13	14	21	14	13	16	11	10
2 category difference	3	4	5	2	2	3	7	10	5	1
3 category difference	1	1	2	4	2	3	2	1	3	1
4 category difference	—	—	1	—	1	—	—	—	1	2
<b>Overestimation (PDW &gt; AAG)</b>										
<b>Total % of participants overestimated by PDW‡</b>	24	27	27	32	22	30	19	18	26	28
1 category difference	19	23	19	27	15	23	13	13	21	22
2 category difference	4	4	6	3	6	4	4	3	3	4
3 category difference	1	1	1	1	—	2	1	2	2	1
4 category difference	—	—	—	1	1	1	—	—	—	—
<b>C. Combined Under and Overestimation (PDW ≠ AAG by 1 category)</b>										
<b>TOTAL DISAGREEMENT</b>	<b>47</b>	<b>57</b>	<b>48</b>	<b>52</b>	<b>48</b>	<b>50</b>	<b>40</b>	<b>46</b>	<b>47</b>	<b>43</b>
Disagreement up to ±1% (difference of up to ±25 steps)	—	—	—	—	—	—	—	—	—	—
Disagreement up to ±5% (difference of up to ±125 steps)	—	—	—	1	—	—	—	—	—	—
Disagreement up to ±10% (difference of up to ±250 steps)	—	—	—	1	—	—	1	—	1	—
Disagreement > ±10% (difference > ±250 steps)	47	57	48	51	48	50	39	46	46	43

All cells are percentages of participants; \*each category represents 2,500 steps (sedentary: < 5,000 steps/day; low active: 5,000-7,499 steps/day; somewhat active: 7,500-9,999 steps/day; active: 10,000-12,499 steps/day; and highly active: ≥ 12,500 steps/day); † total percentages of participants where PDW < AAG by at least one category; ‡ total percentages of participants where PDW > AAG by at least one category.

Table 4 depicts the levels of agreement and disagreement between the PA categories (based on step counts) generated from the data recorded by the two MS (pedometer and accelerometer). Each PA category represented 2,500 steps.

*Agreement of PA categories generated by each MS:* section (A) of Table 4 shows that there was fair total agreement between the two devices (range: 43–60% total agreement) in classifying each participant's step counts into PA categories as recorded by each MS for the same participant.

*Disagreement of PA categories generated by each MS:* when the PA categories generated from the step counts recorded by the two MS were different (*i.e.*, disagreed), we further assessed the direction and size of the disagreement *i.e.*, whether the pedometer generated an underestimation or overestimation in comparison to the accelerometer. Section (B) of Table 4 shows these disagreements (under- and overestimations) grouped by the number of categories difference between the two MS. For both the under- and overestimations, the most frequent disagreements were those where the differences comprised one PA category, followed by disagreements where the differences comprised two categories. It was reassuring to find that disagreements of three PA categories were relatively less common, comprising 1–4% of all cases; and disagreements of four categories were very few (range: 1–2% of all cases). No distinct pattern was observed in the total % of participants underestimated or overestimated across a particular time period or by workday or weekend. However the greatest disagreement (combined under- and overestimations) for the time periods was for the 5-day monitoring (57%); for the individual days, the greatest disagreement was on Monday (52%) followed by Wednesday (50%).

*Disagreement of more than one category of PA:* when the disagreement between the pedometer and accelerometer was of >1 PA category, we undertook no further analysis as such disagreements were deemed important because they have practical implications.

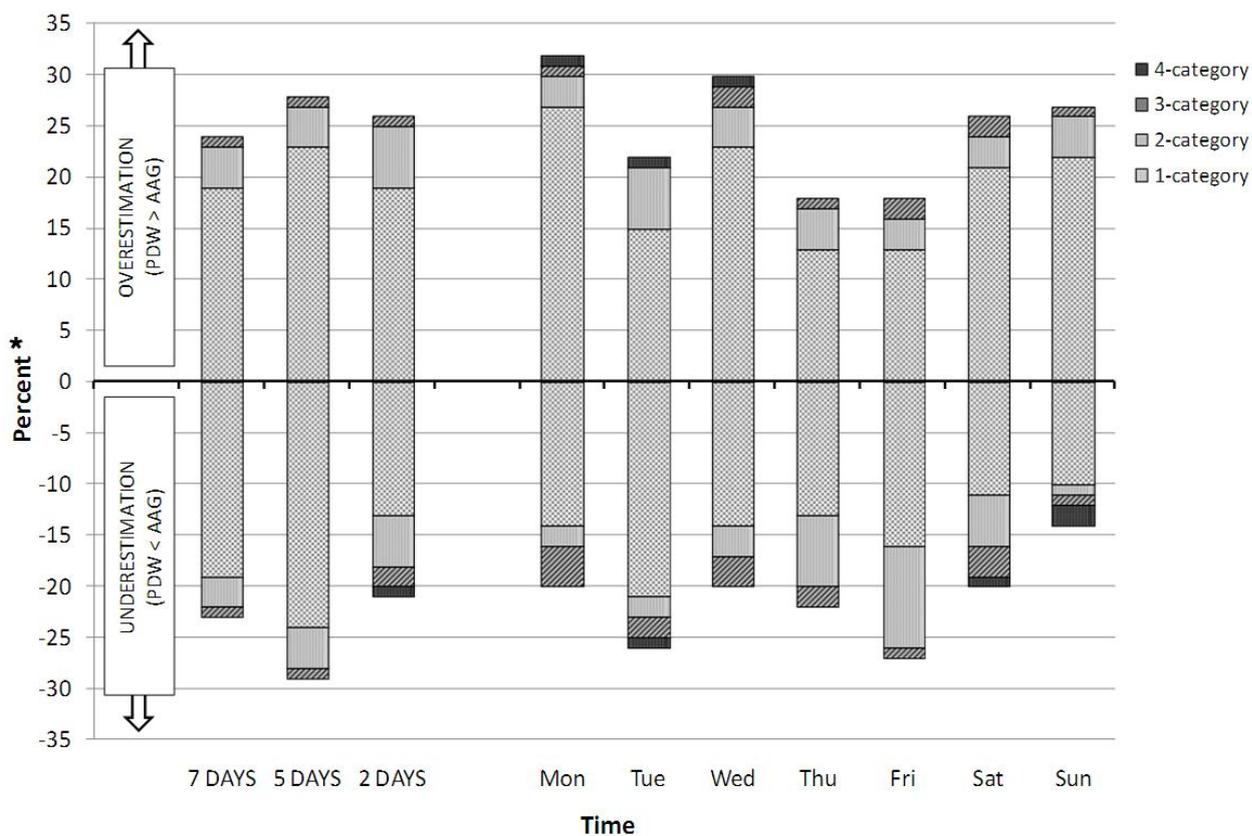
*Disagreement of only one category of PA:* when the disagreement between the two MS was only by 1 PA category, in order to quantify the size of any given extent of disagreement (regardless of whether it was an underestimation or overestimation), we further sub-divided only those disagreements of one PA category into four sub-groups based on the size of the differences in the number of steps recorded by each of the two devices. As each PA category comprised 2,500 steps, four 'disagreement groups' were constructed (Section C of Table 4): disagreement up to  $\pm 1\%$  of the size of the category (*i.e.*, difference of up to  $\pm 25$  steps); disagreement up to  $\pm 5\%$  (*i.e.*, difference of up to  $\pm 125$  steps); disagreement up to  $\pm 10\%$  (*i.e.*, difference of up to  $\pm 250$  steps); and, disagreement  $> \pm 10\%$  (difference  $> \pm 250$  steps).

This procedure enabled us to 'gauge' when a given 1-category disagreement was due to two recorded step counts (pedometer and accelerometer) that were close to but on the opposite sides of the border of the given PA category *i.e.*, when the differences were relatively small—'borderline cases' *e.g.*,  $\pm 1\%$  (difference of  $\pm 25$  steps) or  $\pm 5\%$  (difference of  $\pm 125$  steps); or alternatively, when the differences were larger—'non-borderline cases' *e.g.*, up to  $\pm 10\%$  (difference of  $\pm 250$  steps), and disagreement  $> \pm 10\%$  (difference  $> \pm 250$  steps). Such distinction is important as the existence of many of such 'borderline cases' could lead to misinterpretation (a further overestimation) of the total disagreement values.

Section (C) of Table 4 details the analysis of all 1-category disagreements *i.e.*, differences that could have been due to the step values of each of the two MS falling close to either side of the border

of a given PA category. The findings suggested only a few ‘borderline cases’: none of the sample had differences that were up to  $\pm 1\%$  (*i.e.*, up to  $\pm 25$  steps), and only 1% of the sample showed differences comprising up to  $\pm 5\%$  disagreement category size (*i.e.*, up to  $\pm 125$  steps). Section C also depicts that about 3% of the sample displayed ‘non-borderline cases’ (those comprising up to  $\pm 10\%$  disagreement *i.e.*, up to  $\pm 250$  steps). The highest frequency of disagreements were those where there were  $> \pm 10\%$  disagreement ( $> \pm 250$  steps respectively). For instance 47% of all participants were classified differently by the two devices in the 7-days period, and for all these 47% of participants, the disagreement was  $> \pm 250$  steps (last row of the Table 4). Similarly 57% of all participants were classified differently by the two devices in the 5-days period, and for all these 57% of participants the disagreement was  $> \pm 250$  steps. Across the individual days of monitoring, the disagreements of  $> \pm 250$  steps ranged from 39% (Thursday) to 48% (Monday). This suggested that most of the disagreements (in PA categories generated from the step counts recorded by each of the two MS) were due to ‘non-borderline cases’.

**Figure 1.** Direction and amount of disagreement of participants’ PA categorization based on pedometer (PDW) when compared to accelerometer (AAG) step counts (N = 135).



Note: 1-category: inter-device difference within the range of one PA category ( $\pm 2,500$  steps);  
 2-category: inter-device difference within the range of two categories ( $\pm 5,000$  steps);  
 3-category: inter-device difference within the range of three PA categories ( $\pm 7,500$  steps);  
 4-category: inter-device difference within the range four PA categories ( $\pm 10,000$  steps); \* percent of participants that were over or under estimated by PDW.

Figure 1 illustrates the extents and directions (over- and underestimations) of disagreement between the pedometer and accelerometer when the numbers of steps recorded by each were translated into participants' PA categories. The 7-day period (first column in Figure 1) depicts relatively equal ratios of over- and underestimation of the pedometer as regards the percentages of total disagreement (*i.e.*, all the categories of disagreement combined). The 7-day period also depicts relatively equal ratios of over- and underestimation of the pedometer as regards the percentages of disagreement in the *individual* PA categories. These near-equal ratios of over- and underestimations of the pedometer when compared with the accelerometer give the first column of Figure 1 a 'mirror image' appearance on both sides of the x-axis (the column's height and composition are nearly the same in both directions). However, the 7-day period represented a 'summary' of all the days of the week, and a detailed view of the individual days revealed some discrepancies when compared with the 7-day period. Indeed across the individual days, we noted three points: the unequal lengths of the bars on both sides of the x-axis (unequal *total* disagreement regardless of whether it is was under- or overestimation); there was not a very clear 'mirror image' of near-matching extents of over- and underestimations (unequal lengths of the overestimations and underestimations bars for any individual day); and within the over- or underestimations, there was unequal ratio of the difference categories that comprised the total over- or underestimations (unequal lengths of the subparts that comprise the total overestimation or underestimation bars for any individual day).

#### 4. Discussion

There is increasing interest in objective monitoring of daily PA using electronic MS, e.g., accelerometers and pedometers [25–28]. However the validation of any MS should be undertaken in field settings employing direct observation or doubly-labeled water as criterion measures [29]. Nevertheless, for large studies, these 'gold standard' criteria are expensive and require time. Hence the present study concurrently compared an accelerometer and a pedometer as regards two measurements undertaken by each MS (step counts recorded; and categories of PA calculated). This research contributes to further expand our understanding of the differences in measurements recorded by pedometers and accelerometers (the latter used in this study as the criterion measure). Others have reported the Yamax (SW701) pedometer as an acceptable criterion pedometer [11]. We build on such findings and assessed this Yamax pedometer *vis-a-vis* an accelerometer as a criterion.

As regards the first objective of the study, we appraised the step counts recorded by each of the two MS for the same participant everyday for a week. The step counts of 135 adolescent girls as measured by two MS showed relatively high variability of their PA (IQR range = 4,053 to 6,698 steps per day, Table 1) across the monitored days and the time periods. Mean number of steps recorded by our participants across the week ranged between 6,753–9,414<sub>pedometer</sub> and 6,410–10,043<sub>accelerometer</sub>. When these *step counts* were compared to the guidelines of the 2001–02 President's Challenge Physical Activity and Fitness Awards Program (11,000 steps per day for girls [30]), our sample appeared not to meet this recommended guideline. However, a point to note is that the recommendations of the President's Challenge Physical Activity and Fitness Awards Program [30] were derived from an American population sample. Nevertheless, in the monitored week, regardless the type of MS (pedometer or accelerometer) that recorded the number of steps, the PA levels of these adolescent girls

were mostly ‘somewhat active’ as classified by a highly cited study [1], or as used by Hatano [31]. Interestingly, the PA levels of our adolescent girls were classified as ‘good’ PA level as used by Sigmund and co-researchers [32]. However they [32] used a scale where ‘good’ represented the 4<sup>th</sup> highest level out of six levels.

Regardless of the type of MS, we found significant inter-day differences (across the week) and also significant work- and weekend days differences in the recorded number of steps. Such differences highlight the peculiarity of individual days in relation to the amount of PA achieved by people. These day-to-day differences in PA (more active in weekdays than weekend; least active on Sunday in comparison to the rest of days of the week) are in agreement with others [22,33–38]. People view the weekend as a relaxation time and hence may not push themselves as hard on the weekend [39]. These findings further confirmed that the assessment of PA levels could be strongly dependent on the given day on which PA monitoring is performed. Such features would need to be considered in the research design of studies where free-living PA is evaluated. There seems a lack of consensus of opinions about the day of the week where the highest step counts are usually attained [39,40–43].

As regards the step counts achieved, in relation to the study’s second objective, we compared the number of step recorded by each of the two MS for the same participant in order to detect any inter-device differences on each day over one week. As presented in the results section, the correlations between inter-device differences and the duration of wearing the devices throughout the day were low, suggesting that our findings of the inter-device differences were not influenced by the duration of wearing the devices. Only on two days of the week (Wednesday and Sunday), the numbers of steps recorded by each MS were significantly different (inter-device differences, where the pedometer overestimated the number of steps in comparison to the accelerometer) (Table 1). However, for the whole 7-day period (and for workdays and weekend) there were no significant differences in the number of steps recorded by the two MS, denoting sufficient accordance of both devices. In this sense, we could recommend the use MS (for monitoring the step counts achieved by free living people) when the duration of monitoring is for some longer time period (week, workdays or weekends) rather than on single individual days only.

The mostly positive differences of the measured step counts (differences of devices’ medians:  $Mdn_{AAG} - Mdn_{PDW}$ , Table 1) indicated that the pedometer (PDW) underestimated the number of steps in comparison with the accelerometer (AAG) as a criterion measure. This is in accordance with Le Masurier *et al.* [20] who reported the underestimation of the Yamax SW200 pedometer (in comparison with AAG) under controlled conditions, and the decrease of such underestimation with the increase of the speed of locomotion. We are also in agreement with others [13,44] who have reported similar results: that pedometer underestimated the number of steps achieved when compared with accelerometer.

In our sample, the mostly small effect size (size of the differences, regardless of whether it was under- or overestimation) seems to provide some evidence for the agreement of both devices as step counters. However, we found that the inter-device correlations were not so high (Table 2). Indeed our correlation between the two devices in the 7-day period ( $r_{sp} = 0.65$ ) was mostly lower when compared with other studies that employed exactly the *same* brands of MS (as those that we used) to measure convergent validity of these MS in different populations. For instance, our value (0.65) was lower than in Belgium where  $R$  was 0.73 for 76 children aged 4–5.9 years (monitored for 5days) [16], and also

lower than in Great Britain where  $R$  was 0.86 for 121 older persons > 65 years (monitored for 7 days) [10]. However our correlation findings between the 2 devices were close to that reported in Texas, USA where  $R$  was 0.60 for 78 children 11–15 years. However the Texas study was under controlled conditions rather than for free-living PA [45].

Further, our correlation between the two devices in the 7-day period ( $r_{sp} = 0.65$ ) was also mostly lower when compared with other studies that assessed *different* models of Yamax pedometer (e.g., Yamax SW200, SW500) that the one we used (Yamax SW701) in relation to different brands of accelerometer (other than the Actigraph GT1M which we used). For instance, authors have reported higher convergent validity of Yamax pedometers (SW200, SW500) with various types of accelerometers in adults, e.g., with the Tritrac-Vmag accelerometer,  $R = 0.93$  [46] or CSA accelerometer,  $R = 0.86$  [21, 46]. An explanation for the partial divergence of our findings (our lower correlations between both MS that we used) in relation to other studies could be a consequence of the specific characteristics of PA of the observed population and the fact that we observed a relatively homogeneous sample where there has been reports to suggest that the correlations in homogenous samples could be lower than in non-homogenous samples [47]. Reliability and (more so) validity of frequently used MS depend on the type of PA [21,48] and its intensity [9]. Therefore the preferred type of PA, the levels of various habitual activities (including occupation), and their different intensities (e.g., variance due to gender [49,50] or age [50]) in an observed population have strong relationships to the level of validity of a MS device. Although the correlation between the two MS in our study was relatively lower than others e.g., [10,16], due to the range of afore-mentioned reasons, the MS we assessed can be considered as useful devices to measure the amount of steps of adolescents' free-living PA.

As regards the PA category achieved, in connection with the study's third objective, we classified the step counts recorded by each of the two MS for the same participant into five PA categories, and compared these categories for the same participant in order to detect any inter-device differences in the categories on each day over a period of one week. Although PA categorization based on step counting has been used when assessing the PA levels of free living people [51-54], our ICC values (and their 95% confidence intervals) of the PA categories computed from the step counts of both MS indicated low agreement between the PA categories generated by each of the two devices (ICC = 0.62–0.67). However, ICC values are influenced by many factors (e.g., age; gender; sample size; number of devices that are being compared, *etc.*) [55]. The low ICC between the two devices in our study (0.62–0.67) is further supported by our finding of low (43% to 60%) inter-device agreement of the two MS in categorizing our participants' PA (into 'sedentary', 'low active', 'somewhat active', 'active', and 'highly active' categories). It is difficult for us compare our PA categories findings with others as most studies [14,17,45,56,57] compared the actual number of steps and did not go further to actually compare the PA categories that are generated from these reported number of steps.

When the PA categories generated from the data recorded by the 2 MS were different (*i.e.*, disagreement), we assessed whether the pedometer generated an over- or underestimation in comparison to the accelerometer, and the size of the disagreement. Figure 1 suggested that for the individual days, it was not easy to indicate precisely whether the pedometer generally over- or underestimated the PA categories in comparison to accelerometer. However, for the overall 7-day period, there were comparatively 'balanced' extents of over- and underestimation of the pedometer in

respect to the accelerometer (more symmetrical ‘mirror image’ appearance of the column on both sides of the x-axis). Conversely, the more ‘asymmetric’ columns (in terms of unequal over- and underestimation) were Monday, Wednesday and Sunday. Moreover, Monday and Wednesday were the days with the highest total disagreement; Wednesday and Sunday were the only days with statistically significant inter-device differences in number of steps; and in all three days (Monday, Wednesday and Sunday), the pedometer had the most tendencies to overestimate the PA categories. These findings might suggest some special circumstances/ characteristics of these days (e.g., weather conditions), or alternatively that these differences could be random findings. Further research would confirm or refute such findings.

Our most frequent inter-device disagreements were those where the differences comprised one PA category, followed by disagreements where the differences comprised two categories, with very few cases of disagreements of three and of four categories. Disagreements of *more than* one PA category clearly have practical implications when monitoring adolescents’ PA. Similarly, disagreements of *only* one category could have practical implications if the values generated by the two MS do not lie close to either side of a given category’s boundary *i.e.*, not borderline. We found very few such ‘borderline cases’: none of the sample had small differences (up to  $\pm 25$  steps), and only 1% had differences comprising up to  $\pm 125$  steps. About 3% of our adolescents displayed ‘non-borderline cases’ (those comprising up to  $\pm 250$  steps), and the highest frequency of disagreements were those that comprised  $> \pm 250$  steps. For instance 47% of all participants were classified differently by the two MS in the 7-day period with the magnitude of disagreement being  $> \pm 250$  steps.

Collectively, these results suggested that the disagreement between the two devices in categorizing PA was not mainly due to some values lying close to but on opposite sides of the borders of a given PA category. Rather, most of the disagreement was due to values lying away from and on opposite sides of the borders of a given PA category (non-borderline cases). When this finding is coupled to the fact that we found many high disagreement instances (when the inter-device disagreement was  $>1$  category), this suggested that caution need to be exercised when ‘translating’ pedometer or accelerometer step counts into corresponding PA categories. It was not feasible for us to compare our PA categorization findings as most research only compared step counts rather than PA categories [10,14,16,18,19,57].

Hence our findings suggested that the likelihood of ‘erroneous’ PA categorization (disagreement of categories generated by each MS) for the same participant was not low. Nevertheless, obtaining and assessing information about a person’s PA category is valuable and in many instances such PA category information is combined with other outputs of complex assessments. This proposed that caution needs to be exercised in the interpretation of such PA category data. Our findings also suggested that it could be advantageous to wear two MS devices concurrently in PA research. Concurrent usage of monitors provides more complex and more specific information about the ‘real’ amount of performed PA. It also allows a person’s PA categorization to be made on the basis of steps number obtained from PDW and AAG. Further, wearing both devices facilitates immediate feedback (pedometer) [58-60], and also consequential education-motivated feedback (accelerometer) [61] to the participant. However, the actual price of an accelerometer is several-fold more than the price of a pedometer, a point that may evocate rejection of the concurrent wearing of two MS.

This study has limitations. The data is from girls from four schools in one country. Participants were monitored during autumn, and the structure and amount of PA could be different than during others seasons [39]. Some movement assessment devices have the draw back that their measurements could be impeded when the position of the accelerometer, which should be vertically aligned, is changed in position by soft tissue in the waist region (the place of attachment of the device). Measurements can be affected by such repositioning and might have contributed to the differing results. Further, pedometers are incapable of measuring activity intensity and underestimate the number of steps taken at slower walking speeds, so measurements/ indications of the walking speeds would have been useful.

Although the pedometer is rapidly becoming the preferred objective technique of measuring PA in large populations, previous studies have focused on the quality (e.g., differences in the validity or reliability) of the different brands of pedometers, and have reported non-uniform findings [14]. We found mostly non-significant differences in step counts between the pedometer and accelerometer. Future research could provide more detailed analyses of the factors that might have a negative influence on the validity of pedometer step counting. Further, the immediate feedback that participants received from pedometers could lead to an increased motivation for PA [60], therefore a pedometer's motivational effect size should be described and analyzed.

## **5. Conclusions**

The present study advances our knowledge about MS (pedometer and accelerometer) used for objective assessment of PA. Our sample of Polish adolescent girls PA levels fell mostly in the middle of the PA classifying scale (as “somewhat active” or as “good”), but the variability of recorded PA (steps counts) was relatively high. The girls' PA was not of the same level in all the days of the week; the day-to-day differences and also the differences between weekdays and weekend were statistically significant.

The inter-device comparison of steps achieved showed significant differences on Wednesday and Sunday. There was moderate inter-device correlation for the number of steps achieved on all the days, as well as in the time periods (week, weekdays, and weekend). Using the accelerometer as the criterion, the concurrent validity of the Yamax SW-701 pedometer was considered to be low. However caution needs to be taken when interpreting the findings, as the validity of MS varies with different populations, type of preferred PA, intensity of PA, accessibility of PA and others factors. When these are taken into account, we consider the Yamax SD-701 to generate valuable step count information and to be a useful device to measure the amount of steps achieved by free-living adolescent females.

The total values of the differences in number of steps mostly indicate that the Yamax SW-701 pedometer generally underestimated the step counts in comparison with the ActiGraph GT1M accelerometer, although when there were significant differences between the two MS, the pedometer usually overestimated the step counts.

The likelihood of ‘erroneous’ PA categorization (disagreement of categories generated by each MS) was high. This fact raises two questions: the translation of pedometer and accelerometer step counts into corresponding PA categories; and the ‘credibility’ of the use of five-category classification of PA

(sedentary, low active, somewhat active, active, and highly active). Caution is recommended when using such PA categories in the assessment of people's free-living PA.

MS should be used as PA assessment tools more safely in longer-term periods of monitoring (at least seven days); as in the case of individual-day monitoring, there is the likelihood of significant inter-devices differences that could lead to erroneous interpretation of results. Hence the concurrent usage of both MS could provide more complex information about adolescents' PA, and could assist in immediate feedback to participants (by pedometer) followed by education-motivate feedback (by accelerometer).

## Acknowledgements

The authors acknowledge and thank the schools and participants who took part in this study. The study was supported by a research grant from the Ministry of Education, Youth and Sports of the Czech Republic (No. MSM 6198959221) "Physical Activity and Inactivity of the Inhabitants of the Czech Republic in the Context of Behavioral Changes". We also thank the anonymous reviewers for their constructive comments that helped strengthen the article.

## References

1. Tudor-Locke, C.; Bassett, D.R. How many steps/day are enough? Preliminary pedometer indices for public health. *Sports Med.* **2004**, *34*, 1-8.
2. Sirard, J.R.; Pate, R.R. Physical activity assessment in children and adolescents. *Sports Med.* **2001**, *31*, 439-454.
3. Tudor-Locke, C.; Williams, J.E.; Reis, J.P.; Pluto, D. Utility of pedometers for assessing physical activity: Convergent validity. *Sports Med.* **2002**, *32*, 795-808.
4. Caspersen, C.J.; Powell, K.E.; Christenson, G.M. Physical activity, exercise, and physical fitness: Definitions and distinctions for health-related research. *Public Health Rep.* **1985**, *100*, 126-131.
5. Corder, K.; Ekelund, U.; Steele, R.M.; Wareham, N.J.; Brage, S. Assessment of physical activity in youth. *J. Appl. Physiol.* **2008**, *105*, 977-987.
6. Hands, B.; Parker, H.; Larkin, D. Physical activity measurement methods for young children: A comparative study. *Meas. Phys. Educ. Exerc. Sci.* **2006**, *10*, 203-214.
7. Jürim äe, T.; Jürim äe, J. *Growth, Physical Activity, and Motor Development in Prepubertal Children*; Informa Healthcare: London, UK, 2001; pp. 55-61.
8. Bassett, D.R.; Ainsworth, B.E.; Swartz, A.M.; Strath, S.J.; O'Brien, W.L.; King, G.A. Validity of four motion sensors in measuring moderate intensity physical activity. *Med. Sci. Sport. Exerc.* **2000**, *32*, S471-S480.
9. Crouter, S.E.; Schneider, P.L.; Karabulut, M.; Bassett, D.R. Validity of 10 electronic pedometers for measuring steps, distance, and energy cost. *Med. Sci. Sport. Exerc.* **2003**, *35*, 1455-1460.
10. Harris, T.J.; Owen, C.G.; Victor, C.R.; Adams, R.; Ekelund, U.; Cook, D.G. A comparison of questionnaire, accelerometer, and pedometer: Measures in older people. *Med. Sci. Sport. Exerc.* **2009**, *41*, 1392-1402.

11. Schneider, P.L.; Crouter, S.E.; Bassett, D.R. Pedometer measures of free-living physical activity: Comparison of 13 models. *Med. Sci. Sport. Exer.* **2004**, *36*, 331-335.
12. Felton, G.M.; Tudor-Locke, C.; Burkett, L. Reliability of pedometer-determined free-living physical activity data in college women. *Res. Q. Exercise Sport* **2006**, *77*, 304-308.
13. Schneider, P.L.; Crouter, S.E.; Lukajic, O.; Bassett, D.R. Accuracy and reliability of 10 pedometers for measuring steps over a 400-m Walk. *Med. Sci. Sport. Exer.* **2003**, *35*, 1779-1784.
14. De Vries, S.I.; Van Hirtum, H.W.J.E.M.; Bakker, I.; Hopman-Rock, M.; Hirasings, R.A.; Van Mechelen, W. Validity and reproducibility of motion sensors in youth: a systematic update. *Med. Sci. Sport. Exer.* **2009**, *41*, 818-827.
15. Plasqui, G.; Joosen, A.M.; Kester, A.D.; Goris, A.H.; Westerterp, K.R. Measuring free-living energy expenditure and physical activity with triaxial accelerometry. *Obes. Res.* **2005**, *13*, 1363-1369.
16. Cardon, G.; De Bourdeaudhuij, I. Comparison of pedometer and accelerometer measures of physical activity in preschool children. *Pediatr. Exerc. Sci.* **2007**, *19*, 205-214.
17. Rowlands, A.V.; Eston, R.G. The measurement and interpretation of children's physical activity. *J. Sport Sci. Med.* **2007**, *6*, 270-276.
18. Corder, K.; Brage, S.; Ekelund, U. Accelerometers and pedometers: Methodology and clinical application. *Curr. Opin. Clin. Nutr.* **2007**, *10*, 597-603.
19. Behrens, T.K.; Dinger, M.K. Accelerometer and pedometer step comparisons in free-living adults: 1383: Board #122. *Med. Sci. Sport. Exer.* **2008**, *40*, S199.
20. Le Masurier, G.C.; Lee, S.M.; Tudor-Locke, C. Motion sensor accuracy under controlled and free-living conditions. *Med. Sci. Sport. Exer.* **2004**, *36*, 905-910.
21. Tudor-Locke, C.; Ainsworth, B.E.; Thompson, R.W.; Matthews, C.E. Comparison of pedometer and accelerometer measures of free-living physical activity. *Med. Sci. Sport. Exer.* **2002**, *34*, 2045-2051.
22. Groffik, D.; Frömel, K.; Pelclová J. Pedometer as a method for modification of physical activity in students. *J. Hum. Kinet.* **2008**, *20*, 131-138.
23. Chytil, J. *Vyhodnocení pohybové aktivity v týdenním režimu*, 2006; Software centrum: Olomouc, Czech Republic, 2006.
24. Cohen, J. *Statistical Power Analysis for Behavioral Science*, 2nd ed.; Erlbaum: Hillsdale, IL, USA, 1988; p. 101.
25. Westerterp, K.R. Physical activity assessment with accelerometers. *Int. J. Obes.* **1999**, *23*, S45-S49.
26. Rowlands, A.V.; Eston, R.G.; Ingledeu, D.K. Comparison of accelerometer and pedometer measures of physical activity in boys and girls, ages 8-10 years. *Sports Med.* **1997**, *24*, 258-272.
27. Tudor-Locke, C.; Myers, A.M. Challenges and opportunities for measuring physical activity in sedentary adults. *Sports Med.* **2001**, *31*, 91-100.
28. Tudor-Locke, C.; Myers, A.M. Methodological considerations for researchers and practitioners using pedometers to measure physical (ambulatory) activity. *Res. Q. Exercise Sport* **2001**, *72*, 1-12.
29. Melanson, E.L.J.; Freedson, P.S. Validity of the Computer Science and Applications, Inc. (CCA) activity monitor. *Med. Sci. Sport. Exer.* **1995**, *27*, 934-940.

30. Corben, C.B.; Pangrazi, R.P. Physical activity for children: Current patterns and guidelines. *President's Council on Physical Fitness & Sports Research Digest* **2004**, *5*, 1-5,8.
31. Tudor-Locke, C.; Hatano, Y.; Pangrazi, R.P.; Kang, M. Revisiting “how many steps are enough?”. *Med. Sci. Sport. Exer.* **2008**, *40*, S537-S543.
32. Sigmund, E.; Frömel, K.; Neuls, F. Physical activity of youth: Evaluation guidelines from the viewpoint of health support. *Acta Univ. Palacki. Olomuc. Gymn.* **2005**, *35*, 59-68.
33. Dinger, M.K.; Behrens, T.K. Accelerometer-determined physical activity of free-living college students. *Med. Sci. Sport. Exer.* **2006**, *38*, 774-779.
34. Duncan, J.S.; Schofield, G.; Duncan, E.K. Pedometer-determined physical activity and body composition in New Zealand children. *Med. Sci. Sport. Exer.* **2006**, *38*, 1402-1409.
35. Jones, G.R.; Tudor-Locke, C.; Paterson, D.H. Measuring physical activity in older adult exercisers with pedometers: Importance of assessing weekdays and weekends. *Med. Sci. Sport. Exer.* **2001**, *33*, S252.
36. Peiró-Velert, C.; Devés-Devés, J.; Beltrán-Carrillo, V.J.; Fox, K.R. Variability of Spanish adolescents' physical activity patterns by seasonality, day of the week and demographic factors. *Eur. J. Spor Sci.* **2008**, *8*, 163-171.
37. Tudor-Locke, C.; Ham, S.A.; Macera, C.A.; Ainsworth, B.E.; Kirtland, K.A.; Reis, J.P.; Kimsey, C.D.J. Descriptive epidemiology of pedometer-determined physical activity. *Med. Sci. Sport. Exer.* **2004**, *36*, 1567-1573.
38. Tudor-Locke, C.; Hart, T.L.; Washington, T.L. Expected values for pedometer-determined physical activity in older populations. *Int. J. Behav. Nutr. Phy.* **2009**, *6*, 1-6.
39. Tudor-Locke, C.; Bassett, D.R.; Swartz, A.M.; Strath, S.J.; Parr, B.B.; Reis, J.P.; Dubose, K.D.; Ainsworth, B.E. A preliminary study of one year of pedometer self-monitoring. *Ann. Behav. Med.* **2004**, *28*, 158-162.
40. Chmelík, F.; Frömel, K.; Křen, F.; Stelzer, J.; Engelová, L.; Kudláček, M.; Mitáš, J. The verification of the usability of the online Indares.com system in collecting data on physical activity—pilot study. *Acta Univ. Palacki. Olomuc. Gymn.* **2008**, *38*, 59-66.
41. Gretebeck, R.J.; Montoye, H.J. Variability of some objective measures of physical activity. *Med. Sci. Sport. Exer.* **1992**, *24*, 1167-1172.
42. Michaud, P.A.; Cauderay, M.; Narring, F.; Schutz, Y. Assessment of physical activity with a pedometer and its relationship with VO<sub>2</sub>max among adolescents in Switzerland. *Soz. Präventiv. Med.* **2002**, *47*, 107-115.
43. Vašíčková, J.; Pelclová, J.; Frömel, K.; Chmelík, F.; Pelcl, M. Variability of year-round physical activity in high school girls—pilot study. *Tělesná kultura* **2008**, *31*, 102-108.
44. Bassett, D.R.; Mahar, M.; Rowe, D.A.; Morrow, J.R. Walking and measurement. *Med. Sci. Sport. Exer.* **2008**, *40*, S529-S536.
45. Jago, R.; Watson, K.; Baranowski, T.; Zakeri, I.; Yoo, S.; Baranowski, J.; Conry, K. Pedometer reliability, validity and daily activity targets among 10- to 15-year-old boys. *J. Sport. Sci.* **2006**, *24*, 241-251.
46. Leenders, N.Y.J.M.; Sherman, W.M.; Nagaraja, H.N. Comparisons of four methods of estimating physical activity in adult women. *Med. Sci. Sport. Exer.* **2000**, *32*, 1320-1326.

47. Hopkins, W.G.; Hawley, J.A.; Burke, L.M. Design and analysis of research on sport performance enhancement. *Med. Sci. Sport. Exer.* **1999**, *31*, 472-485.
48. Andre, D.; Wolf, D.L. Recent advance in free-living physical activity monitoring: A review. *J. Diabetes Sci. Technol.* **2007**, *1*, 760-768.
49. Rowlands, A.V.; Eston, R.G. Comparison of accelerometer and pedometer measures of physical activity in boys and girls, ages 8–10 years. *Res. Q. Exercise Sport* **2005**, *76*, 251-257.
50. Trost, S.G.; Pate, R.R.; Sallis, J.F.; Freedson, P.S.; Taylor, W.C.; Dowda, M.; Sirard, J. Age and gender differences in objectively measured physical activity in youth. *Med. Sci. Sport. Exer.* **2002**, *34*, 350-355.
51. Gilson, N.D.; Puig-Ribera, A.; McKenna, J.; Brown, W.J.; Burton, N.W.; Cooke, C.B. Do walking strategies to increase physical activity reduce reported sitting in workplaces: A randomized control trial. *Int. J. Behav. Nutr. Phy.* **2009**, *6*, 43-49.
52. Sigmundová D.; Sigmund, E.; Chmelk, F. Association between the neighborhood and daily steps in Czech regional towns. *Tělesná kultura* **2009**, *32*, 110-124.
53. Whitt-Glover, M.C.; Hogan, P.E.; Lang, W.; Heil, D.P. Pilot Study of a Faith-Based Physical Activity Program Among Sedentary Blacks. *Prev. Chronic. Dis.* **2008**, *5*, 1-9.
54. Wyatt, H.R.; Peters, J.C.; Reed, G.W.; Barry, M.; Hill, J.O. A Colorado Statewide Survey of Walking and Its Relation to Excessive Weight. *Med. Sci. Sport. Exer.* **2005**, *37*, 724-730.
55. Baumgartner, T.A.; Jackson, A.S.; Mahar, M.T.; Rowe, D.A. *Measurement for Evaluation in Physical Education and Exercise Science*; 7th ed.; McGraw-Hill Companies: Boston, MA, USA, 2003; pp. 31-144.
56. Macfarlane, D.J.; Lee, C.C.Y.; Ho, E.Y.K.; Chan, K.L.; Chan, D. Convergent validity of six methods to assess physical activity in daily life. *J. Appl. Physiol.* **2006**, *101*, 1328-1334.
57. McClain, J.J.; Tudor-Locke, C. Objective monitoring of physical activity in children: considerations for instrument selection. *J. Sci. Med. Sport* **2009**, *12*, 526-533.
58. Bravata, D.M.; Smith-Spangler, C.; Sundaram, V.; Gienger, A.L.; Lin, N.; Lewis, R.; Stave, C.D.; Olkin, I.; Sirard, J.R. Using pedometers to increase physical activity and improve health: A systematic review. *J. Amer. Med. Assoc.* **2007**, *298*, 2296-2304.
59. Kang, M.; Marshall, S.; Barreira, T.; Lee, J. Effect of pedometer-based physical activity interventions: A meta-analysis. *Res. Q. Exercise Sport* **2009**, *80*, 648-655.
60. Tudor-Locke, C. Taking steps toward increased physical activity: Using pedometers to measure and motivate. *President's Council on Physical Fitness & Sports Research Digest* **2002**, *3*, 1-6,8.
61. Roemmich, J.N.; Gurgol, C.M.; Epstein, L.H. Open-loop feedback increases physical activity of youth. *Med. Sci. Sport. Exer.* **2004**, *36*, 668-673.