Validity of the Bouchard activity diary in Spanish adolescents

David Martínez-Gómez[1,2], Julia Wärnberg[1,3], Gregory J Welk[4], Michael Sjöström[3], Oscar L Veiga[2] and Ascension Marcos[1,*]

[1]Immunonutrition Research Group, Department of Metabolism and Nutrition, Instituto del Frio, Institute of Food Science, Technology and Nutrition (ICTAN), Spanish National Research Council (CSIC), Calle José Antonio Novais 10, E-28040 Madrid, Spain: [2]Department of Physical Education, Sport and Human Movement, Facultad de Formación del Profesorado y Educación, Universidad Autonoma de Madrid, Madrid, Spain: [3]Unit for Preventive Nutrition, Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, Sweden: [4]Department of Kinesiology, Iowa State University, Ames, IA, USA

Abstract

Objective: Valid and simple instruments to assess physical activity (PA) in specific populations are required for health-related research. The aim of the present study was to assess the validity of the Bouchard activity diary (AD) in Spanish adolescents using an activity monitor to compare total PA and moderate-to-vigorous PA (MVPA) obtained by both instruments.

Design: Sixty-one Spanish adolescents, aged 12–16 years, completed the Bouchard AD and wore the ActiGraph activity monitor for three consecutive days. Validity was assessed with the Spearman correlation coefficient (ρ), the Bland–Altman method and the κ coefficient.

Results: Thirty-seven adolescents were included in the final analysis. Correlations between the activity monitor and the AD administered over the three days (Thursday–Saturday) were moderate ($\rho = 0.33$–0.35, $P < 0.05$) or non-significant for total PA and moderate ($\rho = 0.36$, $P < 0.05$) for MVPA. Correlations between the two methods were progressively lower for each subsequent day of testing, for both total PA and MVPA. The Bland–Altman plot illustrated that the Bouchard AD overestimated MVPA (mean difference -32.05 (SD 74.56) min; 95% limits of agreement 109.61, -173.31 min). Agreements for classification into MVPA tertiles and accordance with the international recommendations of MVPA were fair and moderate, respectively, for the 3 d means.

Conclusions: The Bouchard AD has reasonable validity to assess total PA and MVPA in Spanish adolescents. The results show lower levels of agreement on the third day but it is not clear if this is due to design features (weekday *v.* weekend) or to participant compliance with the survey or the activity monitoring protocol.

The prevalence of overweight and obesity in the Spanish adolescent population is among the highest reported in European countries[1]. Findings from the AVENA (Alimentación y Valoración del Estado Nutricional en Adolescentes; Food and Assessment of the Nutritional Status of Spanish Adolescents) study showed that the prevalence of overweight (including obesity) and obesity was 26% and 6%, respectively, in adolescent males and 19% and 3%, respectively, in adolescent females[2]. Overweight and obesity in childhood and adolescence increase the risk of being overweight or obese in adulthood. National preventive strategies and health policies are clearly needed to invert the increasing rates of overweight and obesity in youth[3–5].

Physical activity (PA) and dietary patterns have been considered as critical behaviours to prevent obesity and co-morbidities[6]. Therefore, in an effort to reduce these conditions, the WHO developed the Global Strategy on Diet, Physical Activity and Health[7]. One of the objectives of the Global Strategy is to develop PA assessment tools that can be used to make more effective international comparisons[7]. There are currently few studies in the scientific literature that have assessed PA in large samples of Spanish adolescents[8–10]. Thus, there is a specific need to develop valid cross-national instruments to assess PA in Spanish adolescents. While activity monitors are widely used, they have limitations for large-scale research applications. Self-report instruments provide a cost-effective strategy for physical activity research while providing valuable contextual information[11]. The activity diary (AD) or activity log is a widely used instrument to assess PA in adults[12,13] and adolescents[14–16]. Although

© The Authors 2009

the AD is categorized as a subjective instrument[11] to assess PA, it has occasionally been considered an objective instrument[17]. As with all assessment techniques, the use of AD has advantages and disadvantages[18]. The main advantages of AD are their low cost and ease of administration to large groups. Disadvantages include challenges with recall, subjectivity and the burden on participants. Challenges associated with using self-report instruments in children are compounded, so that outcomes in this population must often be used with caution[11].

Recent studies in adolescents have found different associations with health outcomes for total PA and PA at different intensities[19,20]. International recommendations for PA for children and adolescents[21] have emphasized moderate-to-vigorous PA (MVPA), so it is important for instruments to be able to capture both total PA and PA at different intensities. The Bouchard AD[22] satisfies this requirement and offers considerable potential as a culturally adaptable instrument for assessing energy expenditure and PA in different populations. Studies have demonstrated the Bouchard AD to have good agreement against doubly labelled water[23] and heart-rate monitors[24] in adolescents, so it can be used with adults as well as with youths. The Bouchard AD[22] is administered during three consecutive days (at least one weekend day) and the PA type, frequency, duration and intensity can be assessed. Objective instruments (e.g. heart-rate monitors, pedometers and activity monitors) are frequently used to validate subjective self-report measures, but activity monitors offer a number of advantages for this type of research. Activity monitors have been used in numerous validity studies in both children/adolescents[25–27] and adults[28–30].

The purpose of the present study was to assess the validity of a 3 d Bouchard AD in a sample of adolescent males and females from Spain. Data from a temporally matched activity monitor were used to assess total levels of PA and time spent in MVPA, and these values were directly compared with the AD. Validity was examined for the full 3 d period as well as separately by day in order to evaluate agreement across three consecutive days.

Methods

Subjects and design
A sample of sixty-one healthy Spanish adolescents (thirty-one males and thirty females) aged 12–16 years was recruited from a high school in the region of Madrid. Before participating in the study, all adolescents were informed of the nature of the study and their parents gave signed written consent. Testing was conducted in accordance with the ethical standards established in the 1961 Declaration of Helsinki (as revised in Hong Kong in 1989 and in Edinburgh in 2000). The study was approved by the Ethics Committee of Puerta de Hierro Hospital (Madrid, Spain).

The sample was divided into five groups (ten to fifteen volunteers per group) and each group was assessed for 1 week with the same procedures. Anthropometric measurements were taken on Wednesdays. Height and weight were measured by standard procedures. BMI was calculated as weight divided by the square of height (kg/m^2). Blood pressure was measured using a digital automatic blood pressure monitor (Omron M6; Omron Health Care Co., Ltd, Kyoto, Japan). The same day, during a physical education class, all participants were provided with instructions about how to complete the AD and how to use the activity monitor. Participants were specifically instructed not to use the activity monitor during showers and water sports, and how to wear the monitor correctly during the daytime. Participants completed the AD for three consecutive days (Thursday, Friday and Saturday) and concurrently wore the activity monitor. The AD and activity monitor were returned the next week, on Tuesday. Total data collection for the study lasted 8 weeks.

Activity diary
The Bouchard 3 d AD[22] consists of a grid that divides the day (24 h) into ninety-six periods of 15 min. Participants categorize the intensity of the primary activity performed in each 15 min block using an intensity scale ranging from 1 to 9. Each numeric activity code refers to a specific level of energy expenditure and can be converted into metabolic ratios of expended energy (MET). The original study by Bouchard *et al.*[22] validated this AD in children and adults using the sum of skinfold thicknesses and percentage of body fat as PA proxy measures. Subsequent studies have validated this AD in adolescents against doubly labelled water[23] and compared it with heart-rate monitoring[24]. These three validation studies[22–24] used different MET values for the intensity categories in their comparison. Table 1 shows MET values used in these three studies. The MET values used in the studies by Bouchard *et al.*[22] and Ekelund *et al.*[24] are similar. However, the MET values used by Bratteby *et al.*[23] are higher in categories 7–9 than in the other two studies. Therefore, in the present study, total PA obtained by the AD (MET·min/d) was calculated using both Bouchard *et al.*'s[22] and Bratteby *et al.*'s[23] MET values for the nine activity codes assigned to each of the 15 min blocks. Total PA (MET·min/d) was calculated as the daily amount of time spent in each period multiplied by the corresponding MET value (e.g. MET·min/d in category 1 = number of periods × 15 min × MET value).

Time spent in MVPA for each day (and the 3 d mean) was also calculated using minutes in categories 7–9 rated in the AD. Ekelund *et al.*[24] used categories 7–9 to compare time spent in MVPA by AD and heart monitors, finding no significant differences between both methods. Only subjects with three completed days (without any gap in the ninety-six blocks) were included in the final sample of that study.

Table 1 MET values used in validation studies of the Bouchard activity diary in adolescents

Intensity category	Examples of activities	MET values		
		Bouchard et al.[22]	Bratteby et al.[23]	Ekelund et al.[24]
1	Sleeping, resting in bed	1·0	0·95	1·0
2	Sitting, eating, writing, etc.	1·5	1·5	1·5
3	Standing, washing, combing hair, etc.	2·3	2·0	2·3
4	Walking indoors, light home activities, etc.	2·8	2·8	2·8
5	Walking outdoors, light manual work, etc.	3·3	3·3	3·5
6	Leisure activities, sport and manual work of low intensity	4·8	4·4	4·5
7	Leisure activities, sport and manual work of moderate intensity	5·6	6·5	5·5
8	Leisure activities, sport and manual work of high intensity	6·0	10·0	6·0
9	Sport activities and work of very high to maximal intensity	7·8	15·0	8·0

MET, metabolic equivalents.

Activity monitor

The ActiGraph GT1M (ActiGraph™ LLC, Fort Walton Beach, FL, USA) activity monitor was used to provide objective information about PA. The ActiGraph (previously known as the MTI and the CSA) is a small and lightweight uniaxial accelerometer (3·8 cm × 3·7 cm × 1·8 cm, 27 g) designed to detect vertical accelerations ranging in magnitude from 0·05 to 2·00***g*** with a frequency response of 0·25–2·50 Hz. The ActiGraph has a sampling frequency of 10 Hz. This activity monitor has been validated in both laboratory and free-living conditions with children and adolescents[31,32]. The activity monitor has been shown to be significantly associated with energy expenditure in children and adolescents[33], and is considered a useful instrument for estimating participation in moderate and vigorous activity[34].

The activity monitor was set to record data using 60 s epochs and was worn on an elastic belt at the lower back. This location, near to the centre of gravity, has been previously used in epidemiological studies with children and adolescents[19,20]. Data were processed and analysed by a JAVA-based software tool developed to analyse the raw output from this activity monitor (*.dat). The software excludes bouts of ten continuous minutes of zeros from the analysis output, considering these periods as non-wearing time. This method has been used in previous studies to ensure that the data are reflective of actual wearing time[35]. The monitors were worn for an average of 765 (SD 109) min during day 1, 789 (SD 129) min during day 2 and 667 (SD 127) min during day 3. An inclusion criterion for the present study was an activity monitor recording with at least 10 h of data per day. Only subjects with valid data from the three consecutive days (Thursday, Friday and Saturday) were included in the final sample. The outcome variables from the activity monitor were total counts per day (counts/d) and total counts adjusted by valid time per day (counts/min·d) and these values were computed for each day and for the 3 d mean. The amounts of time spent in MVPA for each day and for the 3 d mean were also calculated with ≥2000 counts/min as cut-off point. This cut-off point has been used in epidemiological studies to assess MVPA in children and

adolescents[36] and is considered equivalent to a walking rate of about 3 km/h[36].

Statistical analysis

One-way ANOVA was used to examine gender differences in the physical characteristics of the sample. The non-parametric Spearman correlation coefficient (ρ) was used to compare total PA and MVPA measured by the AD and the ActiGraph activity monitor for each day and the 3 d mean. The Wilcoxon rank test was used to compare mean differences in time spent in MVPA between both methods each day and for the 3 d mean. Agreements in estimates of MVPA between the AD and the activity monitor were assessed by the Bland–Altman method[37]. The 95 % limits of agreement were used for describing the total errors between methods. The observed percentage of agreement and Cohen's κ coefficient were used to evaluate the agreement of both PA instruments in classifying adolescents into tertiles and according to the international recommendations in MVPA for children and adolescents[38,39]. The κ statistic was interpreted according to standard convention: almost perfect agreement ($\kappa = 0.81–1.0$), substantial (0·61–0·80), moderate (0·41–0·60), fair (0·21–0·40) or poor (0·00–0·20)[39]. The analyses were performed using the SPSS for Windows statistical software package version 14·0 (SPSS Inc., Chicago, IL, USA) and the level of significance was set at $P < 0.05$ for all analyses.

Results

From the total sample of sixty-one volunteer adolescents, twenty-four were excluded from the final analysis. Five participants did not return the AD on time and one student had an invalid ActiGraph output with only zeros recorded during the 3 d period. The remaining excluded participants did not comply with the AD or the activity monitor criteria, or did not have a complete day of AD and a valid activity monitor recording in the same day. If AD or activity monitor (or both) failed, the participant was excluded. Hence, fifty adolescents (thirty-one males and nineteen females) complied correctly with the first

day, forty-four (twenty-eight males and sixteen females) complied with the first and second days, and finally only a sub-sample of thirty-seven subjects complied with the 3 d period (twenty-three males and fourteen females) of AD and activity monitoring. The physical characteristics of the final sample are displayed in Table 2. There were no differences between genders in most of the physical characteristic variables ($P > 0.05$); the only significant difference between genders was found in systolic blood pressure with males having higher values than females ($P < 0.001$). Therefore, all analyses were performed with males and females combined to increase statistical power.

The values of total PA assessed with both methods are shown in Table 3. Correlations between total PA and MVPA assessed by the AD and the monitor are shown in Table 4. For total PA, concurrent validity showed moderate to high correlations on day 1 ($\rho = 0.57$–0.68, $P < 0.001$) and day 2 ($\rho = 0.53$–0.62, $P < 0.001$) and

moderate ($\rho = 0.33$–0.42, $P < 0.05$) or non-significant ($P > 0.05$) correlations on day 3 and for the 3 d mean. Higher correlations between methods to assess total PA were found when using adjusted counts as the activity monitor variable. Although correlations of the Bouchard AD using different MET values and the activity monitor variables did not maintain a clear trend, the only non-significant correlations between the two methods were found when using Bouchard's MET values on day 3 and for the 3 d mean.

Correlations for time spent in MVPA were high for day 1 ($\rho = 0.65$, $P < 0.001$) but lower for subsequent days. The 3 d mean correlation between methods to assess MVPA was moderate ($\rho = 0.36$, $P < 0.05$). Differences between time spent in MVPA measured by the AD and the activity monitor are shown in Table 5. Statistically significant differences were found only for day 3 and the 3 d mean. The Bland–Altman plot (Fig. 1) for agreement between time spent in MVPA assessed by the Bouchard AD and the activity monitor showed a systematic negative bias. The mean difference was -32.05 (SD 74.56) min (95 % CI for the bias -56.91, -7.19 min), and the limits of agreement were 109.61, -173.31 min. The plot demonstrates that the Bouchard AD tends to underestimate MVPA in highly inactive subjects and to overestimate time in MVPA in highly active subjects. The distribution of points on the Bland–Altman plot (difference compared with the mean of these methods) exhibited a significant negative association ($\rho = -0.56$, $P < 0.01$).

Validity was also assessed by examining classification agreement for levels of PA. Correspondences in tertiles

Table 2 Descriptive physical characteristics of the sample of Spanish adolescents (n 37)

	Males (n 23)		Females (n 14)	
	Mean	SD	Mean	SD
Age (years)	14·7	1·4	15·1	1·2
Height (m)	1·7	0·1	1·6	0·1
Weight (kg)	61·3	13·9	54·0	8·2
BMI (kg/m^2)	22·0	3·3	20·8	2·6
Basal heart rate (beats/min)	72·9	9·9	78·1	9·7
Systolic blood pressure (mmHg)	130·0	11·0	114·1*	7·0*
Diastolic blood pressure (mmHg)	66·5	6·7	66·4	5·7

Mean value was significantly different from that of males: *$P < 0.001$.

Table 3 Total physical activity measured by the Bouchard activity diary and the activity monitor among a sample of Spanish adolescents (n 37)

	Day 1 (Thursday)		Day 2 (Friday)		Day 3 (Saturday)		3 d mean	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bouchard activity diary								
Bouchard *et al.*'s MET values (MET·min/d)	2779	569	2931	585	2963	672	2891	460
Bratteby *et al.*'s MET values (MET·min/d)	3272	984	3554	1157	3661	1394	3496	861
Activity monitor								
Total counts (counts/d × 10^{-3})	446	210	538	247	446	181	477	154
Adjusted counts (counts/min·d)	586	268	683	300	679	257	649	202

MET, metabolic equivalents.

Table 4 Spearman correlations for total physical activity and moderate-to-vigorous physical activity (MVPA) measured by the Bouchard activity diary and the activity monitor among a sample of Spanish adolescents (n 37)

Activity monitor	Bouchard activity diary	Day 1 (Thursday)	Day 2 (Friday)	Day 3 (Saturday)	3 d mean
Total counts (counts/d × 10^{-3})	Bouchard *et al.*'s MET values (MET·min/d)	0·57***	0·62***	0·30	0·29
	Bratteby *et al.*'s MET values (MET·min/d)	0·59***	0·56***	0·39*	0·35*
Adjusted counts (counts/min·d)	Bouchard *et al.*'s MET values (MET·min/d)	0·68***	0·62***	0·35*	0·33*
	Bratteby *et al.*'s MET values (MET·min/d)	0·68***	0·53***	0·42*	0·35*
MVPA (min)†	MVPA (min)‡	0·65***	0·42**	0·41*	0·36*

MET, metabolic equivalents.
Correlation was statistically significant: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.
†Calculated with the cut-off of >2000 counts/min for MVPA from the ActiGraph activity monitor.
‡Calculated with the intensity categories 7–9 from the Bouchard activity diary.

Table 5 Differences between time spent in moderate-to-vigorous physical activity (MVPA) measured by the Bouchard activity diary and the activity monitor among a sample of Spanish adolescents (*n* 37)

	Bouchard activity diary MVPA (min)†		Activity monitor MVPA (min)‡		
	Mean	SD	Mean	SD	*P* value
Day 1	107·4	91·3	81·9	47·2	0·078
Day 2	113·5	102·1	104·3	58·7	0·729
Day 3	141·1	127·2	79·7	43·6	0·012
3 d mean	120·7	77·1	88·6	36·2	0·027

Differences were analysed by the Wilcoxon test and statistical significance was set at $P < 0.05$.
†Calculated with the intensity categories 7–9 from the Bouchard activity diary.
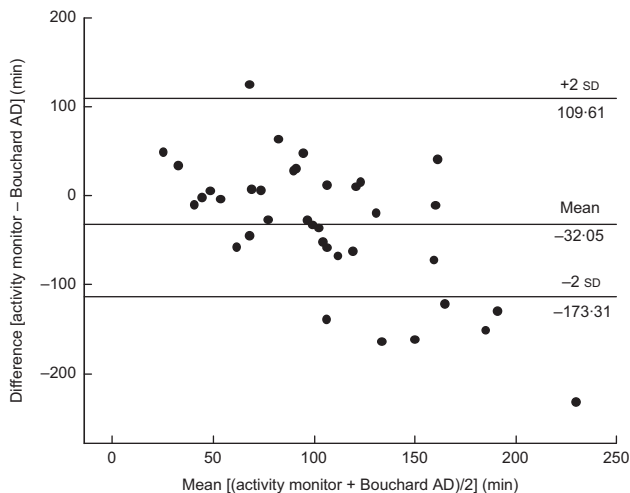‡Calculated with the cut-off of >2000 counts/min for MVPA from the ActiGraph activity monitor.

Fig. 1 Bland–Altman plot for the 3 d mean time spent in moderate-to-vigorous physical activity (MVPA, min) measured by the activity monitor and the Bouchard activity diary (AD; *n* 37). The differences and the means of time spent in MVPA by the two methods were significantly correlated ($\rho = -0.56$, $P < 0.01$)

(low, medium and high) of MVPA and compliance with the guidelines for adolescents of 60 min of MVPA daily[21] are provided in Table 6. Percentage of agreement in tertiles was poor to fair across days and for the 3 d mean. Days 1 and 2 were above 50 % but lower values were observed on day 3 and for the 3 d mean (this effect for the 3 d mean was again influenced predominantly by the values on day 3). Observed percentage of agreement for correct classification according to the international recommendation was less than 70 % only on day 3; the highest value was obtained for the 3 d mean with ~84 % of agreement between methods. The κ statistic was fair on days 1 and 3, moderate on day 2 and for the 3 d mean. The highest κ value was also obtained for the 3 d mean.

Discussion

The present study examined the validity of the Bouchard AD to assess total PA and MVPA in Spanish adolescents using an activity monitor as an objective criterion measure.

The findings show that total PA variables assessed by the ActiGraph over three consecutive days were moderately correlated ($\rho = 0.33$ and 0.35) against total PA assessed by the AD using Bratteby's MET values but not using Bouchard's MET values. A moderate correlation ($\rho = 0.36$) was also found when comparing time spent in MVPA between both methods. Similar results with activity monitors have been reported for other subjective instruments used to assess PA in adolescents[25–27].

The Bland–Altman plot indicates that the Bouchard AD tended to overestimate MVPA in adolescents compared with the activity monitor and that the mean difference in time spent in MVPA between both methods was significant ($P = 0.027$). Seven highly active subjects were outside the limits of agreement. This result may be due to the fact that the activity monitors do not properly detect static activities (e.g. strength or flexibility activities) or activities on the horizontal axis (e.g. cycling, skating) or because the monitor was instructed to be removed in activities with water (e.g. swimming, water polo, skiing). Ekelund *et al.*[24] compared the Bouchard AD against heart-rate monitors in Swedish adolescents and the Bland–Altman plot showed no differences between means for time spent in MVPA. A relevant fact comparing both studies is the mean time spent in MVPA across the three days. The Swedish adolescents used in the validation study spent 38 min while the Spanish adolescents from our study spent an average of 121 min in MVPA (classified as blocks rated from 7 to 9 on the Bouchard AD). Conclusions cannot be drawn when comparing the two study results, but Spanish adolescents might overestimate the time spent in MVPA using the Bouchard AD more than Swedish or else the Spanish sample selected for the present study was more active than the Swedish adolescents in Ekelund *et al.*'s study[24].

Another noteworthy finding of our study was the apparent decrease in validity coefficients over the 3 d period. Previous validation and comparison studies with activity diaries have used the average of several days as a representative weekly PA value[22–24] or used only a 24 h period of administration[40]. The assessments of total PA between the two methods across the days showed moderate to high correlations ($\rho = 0.57$–0.68) on day 1 (Thursday) as well as on day 2 ($\rho = 0.53$–0.62). However,

Table 6 Percentage of agreement between the Bouchard activity diary and the activity monitor in classifying a sample of Spanish adolescents (n 37) according to their time spent in moderate-to-vigorous physical activity (MVPA)

	Tertiles (low, medium, high)		Recommendations for youth†	
	% Agreement on MVPA	κ	% Agreement on MVPA	κ
Day 1	51·35	0·27	70·27	0·34
Day 2	59·46	0·39	72·97	0·41
Day 3	40·54	0·11	67·57	0·30
3 d mean	45·95	0·19	83·78	0·56

†60 min MVPA/d[21].

correlations on day 3 (Saturday) were moderate ($\rho = 0\cdot35$–$0\cdot42$) or not significant using the MET values proposed by Bouchard. These results suggest that the Bouchard AD showed good validity for assessing total PA in adolescents on the first two days but the lower value on the third day affected the overall agreement for the 3 d mean.

The study also examined results using different MET values for the intensity categories 1–9, but it is not possible to determine which is best. A reasonable approach that we would propose is to use Bratteby's[23] MET values in adolescents and Bouchard's[22] MET values in adults since several validation studies have found good agreement in measuring total PA[23,40] in these age ranges. Likewise, the use of Bratteby's MET values in adolescents is more consistent with the new Compendium of Energy Expenditures for Youth[41].

The correlations across days between methods to assess MVPA had a similar trend to that found for total PA. There was a high correlation on day 1 ($\rho = 0\cdot65$) while correlations on days 2 and 3 were decreased slightly ($\rho = 0\cdot42$ and $0\cdot41$, respectively). The time spent in MVPA showed no differences on days 1 and 2 but was significantly different on day 3 and for the average of the three days (see Table 5). Several factors may cause this decline in accuracy on the third day. First, the third day fell on a weekend day (Saturday). During weekend days, adolescents have fewer responsibilities and their activities tend to be more spontaneous and intermittent[42–44]. This fact might make the classification into activity categories 1–9 more difficult or cause a possible excessive use of memory by completing the AD fewer times this day. Second, three days of administration might have caused a loss of motivation and concentration in adolescents for completing the AD. Third, although both instruments are non-invasive techniques, adolescents had to perform all instructions to complete the AD and wear the activity monitor correctly to be included. These increased responsibilities could cause a potential loss of motivation or concentration after the first two days of administration. More research is needed to better understand motivational and behavioural factors that may influence the validity of self-report assessments. Additional work in this area might help researchers to improve the accuracy of the data or to minimize problems with compliance. If the Bouchard AD is used in adolescents over a 3 d period,

some type of external input may help to improve motivation and compliance over the 3 d period. SMS texts, emails or telephone calls may also help to motivate the subject during the period of assessment. These types of external motivation have been used successfully in previous studies with activity monitors[45].

Another aspect examined in the current study was the ability to classify subjects into tertiles and according to the current international PA recommendation for youth[21] (60 min MVPA/d). When classifying into tertiles (day-by-day and for the 3 d mean) only fair agreements were found (3 d mean ~46 %). These results are consistent with previous results in adolescents[24] and adults[46] using the Bouchard AD. Moreover, when using the cut-off of 60 min MVPA/d an agreement of ~84 % ($\kappa = 0\cdot56$) with the 3 d mean was found, this being higher than the agreements found day-by-day. These results show that the Bouchard AD has a reasonable validity to classify subjects in accordance with the current PA recommendation for youth and is fair at classifying into tertiles of total PA.

A similar validation study in adolescents has been performed with an adapted version of the original Previous Day Physical Activity Recall (PDPAR) for adolescents, administered as an AD during 4 d[47]. Moderate correlations ($\rho = 0\cdot41$–$0\cdot42$) were found between concurrent minutes in MVPA obtained by this diary over 4 d and the MTI activity monitor (previous version of the ActiGraph). Once again, the AD overestimated MVPA with respect to the activity monitor ($-38\cdot9$ min and $-86\cdot7$ min in MVPA using different cut-off values by the activity monitor). The authors also discussed the capacity to classify MVPA tertiles and reported an agreement of 45 % ($\kappa = 0\cdot17$). The Bland–Altman plot showed a reasonable agreement between the AD and activity monitor for ≤60 min in MVPA.

Several limitations of our study should be highlighted. First, 40 % of the volunteers were excluded from the analysis reducing the final sample to only thirty-seven adolescents. This reduction or dropout was not exclusively caused by the AD administration (see causes for exclusion in Results section). Second, there is no consensus yet on what cut-off values are better to assess the time spent in different PA intensities using the ActiGraph activity monitor in children. Differences between PA

intensity cut-off values are a fact[48], as are the differences with other PA instruments[47]. Finally, three consecutive days might be insufficient to represent the MVPA of an adolescent's week. The original validity study of the Bouchard AD[22] showed excellent 1-week test–retest reliability to assess total PA for 3 d (intra-class correlation coefficient = 0·96); however, reliability of assessment of minutes in categories 7–9 during 3 d was low (intra-class correlation coefficient = 0·48). In order to get a reliability of 0·80 it would be necessary to include more days of administration. For the adolescent PDPAR diary, 12 d was predicted to be necessary to obtain a reliability of 0·80[47].

In summary, the Bouchard AD showed an acceptable validity for assessing total PA and MVPA in Spanish adolescents. Its use over three consecutive days (two week-days and one weekend day) showed good validity on the first two days but lower validity on the third day. An additional external motivation (SMS, email, telephone calls) during the third day and/or the weekend day might improve the accuracy of the AD in adolescents. Likewise, the findings in the current study and the coherence with the new Compendium of Energy Expenditures for Youth support the use of Bratteby's MET values to assess PA in adolescents by the Bouchard AD.

Acknowledgements

References

1. Flodmark CE, Lissau I, Moreno LA, Pietrobelli A & Widhalm K (2004) New insights into the field of children and adolescents' obesity: the European perspective. *Int J Obes Relat Metab Disord* **28**, 1189–1196.
2. Moreno LA, Mesana MI, Fleta J, Ruiz JR, González-Gross M, Sarría A, Marcos A & Bueno M; AVENA Study Group (2005) Overweight, obesity and body fat composition in Spanish adolescents. The AVENA Study. *Ann Nutr Metab* **49**, 71–76.
3. Aranceta J, Pérez-Rodrigo C, Serra-Majem L, Bellido D, de la Torre ML, Formiguera X & Moreno B (2007) Prevention of overweight and obesity: a Spanish approach. *Public Health Nutr* **10**, 1187–1193.
4. Ballesteros Arribas JM, Dal-Re Saavedra M, Pérez-Farinós N & Villar Villalba C (2007) The Spanish strategy for nutrition, physical activity and the prevention of obesity (NAOS Strategy). *Rev Esp Salud Publica* **81**, 443–449.
5. Singh AS, Mulder C, Twisk JW, van Mechelen W & Chinapaw MJ (2008) Tracking of childhood overweight into adulthood: a systematic review of the literature. *Obes Rev* **9**, 474–488.
6. World Health Organization (2003) *Diet, Nutrition and the Prevention of Chronic Diseases. Report of a Joint WHO/FAO Expert Consultation. WHO Technical Report Series* no. 916. Geneva: WHO.
7. World Health Organization (2004) *Global Strategy on Diet, Physical Activity and Health.* Geneva: WHO.
8. Henríquez Sánchez P, Doreste Alonso J, Laínez Sevillano P, Estévez González MD, Iglesias Valle M, Martín López G, Sosa Iglesias I & Serra Majem L (2008) Prevalence of obesity and overweight in adolescents from Canary Islands, Spain. Relationship with breakfast and physical activity. *Med Clin (Barc)* **130**, 606–610.
9. García-Artero E, Ortega FB, Ruiz JR, Mesa JL, Delgado M, González-Gross M, García-Fuentes M, Vicente-Rodríguez G, Gutiérrez A & Castillo MJ (2007) Lipid and metabolic profiles in adolescents are affected more by physical fitness than physical activity (AVENA study). *Rev Esp Cardiol* **60**, 581–588.
10. Serra-Majem L, Aranceta Bartrina J, Pérez-Rodrigo C, Ribas-Barba L & Delgado-Rubio A (2006) Prevalence and determinants of obesity in Spanish children and young people. *Br J Nutr* **96**, Suppl. 1, S67–S72.
11. Sirard JR & Pate RR (2001) Physical activity assessment in children and adolescents. *Sports Med* **31**, 439–454.
12. Simonen RL, Perusse L, Rankinen T, Rice T, Rao DC & Bouchard C (2002) Familial aggregation of physical activity levels in the Québec Family Study. *Med Sci Sports Exerc* **34**, 1137–1142.
13. Wickel EE & Eisenmann JC (2006) Within- and between-individual variability in estimated energy expenditure and habitual physical activity among young adults. *Eur J Clin Nutr* **60**, 538–544.
14. Eisenmann JC & Wickel EE (2007) Estimated energy expenditure and physical activity patterns of adolescent distance runners. *Int J Sport Nutr Exerc Metab* **17**, 178–188.
15. Jago R, Anderson CB, Baranowski T & Watson K (2005) Adolescent patterns of physical activity differences by gender, day, and time of day. *Am J Prev Med* **28**, 447–452.
16. Hart JW, Bremner AR, Wootton SA & Beattie RM (2005) Measured versus predicted energy expenditure in children with inactive Crohn's disease. *Clin Nutr* **24**, 1047–1055.
17. LaMonte MJ, Ainsworth BE & Tudor-Locke C (2003) Assessment of physical activity and energy expenditure. In *Obesity: Etiology, Assessment, Treatment and Prevention*, pp. 111–140 [RE Andersen, editor]. Champaign, IL: Human Kinetics.
18. Matthews CE (2002) Use of self-report instruments to assess physical activity. In *Physical Activity Assessment in Health Related Research*, pp. 108–121 [GJ Welk, editor]. Champaign, IL: Human Kinetics.
19. Ruiz JR, Rizzo NS, Hurtig-Wennlöf A, Ortega FB, Wärnberg J & Sjöström M (2006) Relations of total physical activity and intensity to fitness and fatness in children: the European Youth Heart Study. *Am J Clin Nutr* **84**, 299–303.
20. Rizzo NS, Ruiz JR, Hurtig-Wennlöf A, Ortega FB & Sjöström M (2007) Relationship of physical activity, fitness, and fatness with clustered metabolic risk in children and adolescents: the European Youth Heart Study. *J Pediatr* **150**, 388–394.

21. Strong WB, Malina RM, Blimkie CJ *et al.* (2005) Evidence based physical activity for school-age youth. *J Pediatr* **146**, 732–737.

22. Bouchard C, Tremblay A, Leblanc C, Lortie G, Savard R & Thériault G (1983) A method to assess energy expenditure in children and adults. *Am J Clin Nutr* **37**, 461–467.

23. Bratteby LE, Sandhagen B, Fan H & Samuelson G (1996) A seven day activity diary for assessment of daily energy expenditure validated by the doubly labeled water method. *Eur J Clin Nutr* **50**, 284–289.

24. Ekelund U, Yngve A & Sjöström M (1999) Total daily energy expenditure and patterns of physical activity in adolescents assessed by two different methods. *Scand J Med Sci Sports* **9**, 257–264.

25. Pate RR, Ross R, Dowda M, Trost SG & Sirard JR (2003) Validation of a 3-day physical activity recall instrument in female youth. *Pediatr Exerc Sci* **15**, 257–265.

26. Janz KF, Lutuchy EM, Wenthe P & Levy SM (2008) Measuring activity in children and adolescents using self-report: PAQ-C and PAQ-A. *Med Sci Sports Exerc* **40**, 767–772.

27. Wong SL, Leatherdale ST & Manske SR (2006) Reliability and validity of a school-based physical activity questionnaire. *Med Sci Sports Exerc* **38**, 1593–1600.

28. Hagströmer M, Oja P & Sjöström M (2006) The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. *Public Health Nutr* **9**, 755–762.

29. Craig CL, Marshall AL, Sjöström M *et al.* (2003) International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* **35**, 1381–1395.

30. Yore MM, Ham SA, Ainsworth BE, Kruger J, Reis JP, Kohl HW 3rd & Macera CA (2007) Reliability and validity of the instrument used in BRFSS to assess physical activity. *Med Sci Sports Exerc* **39**, 1267–1274.

31. Puyau MR, Adolph AL, Vohra FA & Butte NF (2002) Validation and calibration of physical activity monitors in children. *Obes Res* **10**, 150–157.

32. Trost SG, Ward DS, Morread SM, Watson PD, Riner W & Burke J (1998) Validity of the Computer Science and Applications (CSA) activity monitor in children. *Med Sci Sports Exerc* **30**, 629–633.

33. Ekelund U, Sjöström M, Yngve A, Poortvliet E, Nilsson A, Froberg K, Wedderkopp N & Westerterp K (2001) Physical activity assessed by activity monitor and doubly labeled water in children. *Med Sci Sports Exerc* **33**, 275–281.

34. Trost SG, Way R & Okely AD (2006) Predictive validity of three ActiGraph energy expenditure equations for children. *Med Sci Sports Exerc* **38**, 380–387.

35. Andersen LB, Harro M, Sardinha LB, Froberg K, Ekelund U, Brage S & Anderssen SA (2006) Physical activity and clustered cardiovascular risk in children: a cross-sectional study (The European Youth Heart Study). *Lancet* **368**, 299–304.

36. Brage S, Wedderkopp N, Anderson LB & Froberg K (2003) Influence of step frequency on movement intensity predictions with the CSA accelerometer: a field validation study in children. *Pediatr Exerc Sci* **15**, 277–287.

37. Bland JM & Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1**, 307–310.

38. Cohen J (1960) A coefficient of agreement for nominal scales. *Educ Psychol Meas* **20**, 37–46.

39. Landis JR & Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* **33**, 159–174.

40. Wickel EE, Welk GJ & Eisenmann JC (2006) Concurrent validation of the Bouchard diary with an accelerometry-based monitor. *Med Sci Sports Exerc* **38**, 373–379.

41. Ridley K, Ainsworth BE & Olds TS (2008) Development of a compendium of energy expenditures for youth. *Int J Behav Nutr Phys Act* **5**, 45.

42. Nader PR, Bradley RH, Houts RM, McRitchie SL & O'Brien M (2008) Moderate-to-vigorous physical activity from ages 9 to 15 years. *JAMA* **300**, 295–305.

43. Rowlands AV, Pilgrim EL & Eston RG (2008) Patterns of habitual activity across weekdays and weekend days in 9–11-year-old children. *Prev Med* **46**, 317–324.

44. Aires L, Santos R, Silva P, Santos P, Oliveira J, Ribeiro JC, Rego C & Mota J (2007) Daily differences in patterns of physical activity among overweight/obese children engaged in a physical activity program. *Am J Hum Biol* **19**, 871–877.

45. Trost SG, McIver KL & Pate RR (2005) Conducting accelerometer-based activity assessments in field-based research. *Med Sci Sports Exerc* **37**, 11 Suppl., S531–S543.

46. Schmidt MD, Freedson PS & Chasan-Taber L (2003) Estimating physical activity using the CSA accelerometer and a physical activity log. *Med Sci Sports Exerc* **35**, 1605–1611.

47. Anderson CB, Hagströmer M & Yngve A (2005) Validation of the PDPAR as an adolescent diary: effect of accelerometer cut points. *Med Sci Sports Exerc* **37**, 1224–1230.

48. Guinhouya CB, Hubert H, Soubrier S, Vilhelm C, Lemdani M & Durocher A (2006) Moderate-to-vigorous physical activity among children: discrepancies in accelerometry-based cut-off points. *Obesity (Silver Spring)* **14**, 774–777.