LASER WORDS	JOSH	josh˙12418	B	Dispatch: June 15, 2016	Journal: JOSH	CE: Anbarasi.D
	Journal Name	Manuscript No.		Author Received:	No of pages: 9	TS: karthik

RESEARCH ARTICLE

Reliability and Validity of the PAQ-C Questionnaire to Assess Physical Activity in Children

AQ1 JAVIER BENÍTEZ-PORRES, PhD, MS[a] IVÁN LÓPEZ-FERNÁNDEZ, PhD[b] JUAN FRANCISCO RAYA, MS[c] SABRINA ÁLVAREZ CARNERO, BE[d] JOSÉ RAMÓN ALVERO-CRUZ, MD, PhD[e] ELVIS ÁLVAREZ CARNERO, PhD[f]

ABSTRACT

BACKGROUND: Physical activity (PA) assessment by questionnaire is a cornerstone in the field of sport epidemiology studies. The Physical Activity Questionnaire for Children (PAQ-C) has been used widely to assess PA in healthy school populations. The aim of this study was to evaluate the reliability and validity of the PAQ-C questionnaire in Spanish children using triaxial accelerometry as criterion.

METHODS: Eighty-three (N = 46 boys, N = 37 girls) healthy children (age 10.98 ± 1.17 years, body mass index 19.48 ± 3.51 kg/m²) were volunteers and completed the PAQ-C twice and wore an accelerometer for 8 consecutive days. Reliability was analyzed by the intra-class correlation coefficient (ICC) and the internal consistency by the Cronbach's α coefficient. The PAQ-C was compared against total PA and moderate to vigorous PA (MVPA) obtained by accelerometry.

RESULTS: Test-retest reliability showed an ICC = 0.96 for the final score of PAQ-C. Small differences between first and second questionnaire administration were detected. Few and low correlations (rho = 0.228-0.278, all ps < .05) were observed between PAQ-C and accelerometry. The highest correlation was observed for item 9 (rho = 0.311, p < .01).

CONCLUSIONS: PAQ-C had a high reliability but a questionable validity for assessing total PA and MVPA in Spanish children. Therefore, PA measurement in children should not be limited only to self-report measurements.

Keywords: accelerometry; PAQ-C; child health; validation; self-reported physical activity.

Citation: Porres JB, Fernández IL, Raya JF, Carnero SÁ, Cruz JRA, Carnero EÁ. Reliability and Validity of the PAQ-C Questionnaire to Assess Physical Activity in Children. J Sch Health. 2016; 86: 683-691.

Received on November 2, 2014
Accepted on March 24, 2016

Physical activity (PA) has been identified as an important agent in the prevention of chronic diseases such as obesity, cardiovascular diseases, and metabolic syndrome.[1-3] Therefore, to know more precisely the levels of PA during childhood and to identify the impact on health, it is necessary to develop and validate instruments able to assess PA in school populations.

Survey instruments continue to provide useful information in population-based studies of children's PA. The most accurate method for measuring energy expenditure during PA, such as doubly labeled water or indirect calorimetry are complex, time-consuming, and are expensive and impractical procedures when evaluating large populations.[4] Over the last decade, a rise in the technology of accelerometers has allowed

[a]Professor, (benitez@uma.es), Biodynamic and Body Composition Laboratory, Faculty of Education Sciences, University of Malaga, Room 2.01-A - Faculty of Educational Sciences, Bulevar Louis Pasteur, 25, Campus of Teatinos, 29010 Málaga, Spain.

[b]Professor, (ivanl@uma.es), Biodynamic and Body Composition Laboratory, Faculty of Education Sciences, University of Malaga, Room 261 - Faculty of Educational Sciences, Bulevar Louis Pasteur, 25, Campus of Teatinos, 29010 Málaga, Spain.

[c]Teacher of Physical Education, (juanfran_lee@yahoo.es), San Sebastián Primary School, Mexico, S/N, 29650 Mijas, Málaga, Spain.

[d]Teacher of Physical Education, (sabrinaalcar@edu.xunta.es), Julio Gurriarán Primary School, Lugo, 2, 32[illegible] Barco, Galicia, Spain.

[e]Professor (alvero@uma.es), Sport Medicine School, Andalucía Tech, Faculty of Medicine, University of Malaga, Campus of Teatinos S/N, 29071 Málaga, Spain.

[f]Professor, (ecarnero@uma.es), Biodynamic and Body Composition Laboratory, Faculty of Education Sciences, University of Malaga, Room 2.01-A - Faculty of Educational Sciences, Bulevar Louis Pasteur, 25, Campus of Teatinos, 29010 Málaga, Spain.

Address correspondence to Javier Benítez Porres, Professor, (benitez@uma.es), Biodynamic and Body Composition Laboratory, Faculty of Education Sciences, University of Malaga, Room 2.01-A - Faculty of Educational Sciences, Bulevar Louis Pasteur, 25, Campus of Teatinos, 29010 Málaga, Spain.

This work was supported by the Spanish Ministry of Education, Culture and Sport (AP2010-0583); Spanish Ministry of Economy and Competitiveness (DEP2011-30565); and University of Málaga (Campus of International Excellence Andalucía Tech).

us to obtain reliable measurements of the duration of PA, and has provided an indirect, yet reasonably accurate measure of PA in this area. However, this continues to be an expensive method for application in certain settings such as primary schools. Self-report instruments provide a convenient way to assess activity patterns in large populations; however, there are few questionnaires designed to estimate PA in children.

One of the most widely used questionnaires for this age group is the Physical Activity Questionnaire for Children (PAQ-C),[5] though it has never been validated with Spanish children. The PAQ-C is a simple questionnaire that assesses the PA, a child has performed over the last 7 days. The overall result of the test is a score of 1 to 5 points that allows for a graded level of PA performed by each subject. PA measured by PAQ-C has been associated with indicators of adiposity, bone mineral content, heart rate variability, and certain psychological indicators (sports competition, body satisfaction, anxiety).[6-8] The PAQ-C has acceptable reliability and convergent validity.[9] The mean of all items is used to indicate the level of PA. A high score indicates higher levels of PA. Moreover, attempts to obtain ~~cut-off~~ points from PAQ-C final score have been reported for English children.[10] In addition, this questionnaire has adequate validity and reliability in other countries,[9] although the validity and reliability assessed varies by ethnicity and requires additional development to be a useful measure of PA in American and European children from diverse ethnic populations.[11] However, not enough is known about the capacity of PAQ-C to estimate PA measured objectively. Moreover, discrepancies and high variability in children's PA measured by accelerometers around the world have been reported, which may introduce a bias in validation studies.[12,13] So, transcultural validation of PA questionnaires enables us to compare results among countries and a more reliable and valid assessment of PA in each country.

Considering the need for adequate and viable methods of measuring PA in school settings, the aim of this study was to evaluate the reliability and validity of the PAQ-C when applied to Spanish children using the objective measurement of PA by triaxial accelerometry as a reference criterion.

METHODS

Participants

An invitation to participate in the study was sent to all parents who had their children in fourth, fifth, or sixth grade in ~~two~~ different primary schools (Málaga and Orense, Spain). Eighty-three potentially eligible subjects responded, and gave their written informed consent after receiving detailed information about the aims and procedures of the study. Youth with incomplete PA (N = 0) data or technical errors in the instrument (N = 5) were excluded. A final sample of 83 children (46 boys, 37 girls) participated in the reliability study and 78 children (42 boys, 36 girls) cover the criteria for the validity study. There were no differences on age and body mass index (BMI) between the excluded participants and the final sample.

This study is part of ~~two~~ larger studies where the sample size was randomized from all eligible students of each primary school who met the inclusion criteria. Briefly, in the Orense school we had only 30 possible candidates (budget limitation) to participate in a doubly labeled water follow-up study during 4 weeks within 1 year. One hundred-fifty informed consents were distributed among children who were interested in the study. Of those, 25 met the inclusion criteria and only 19 completed the final protocol.

Regarding the Malaga school, all students were called to participate in a randomized control trial study, and 100 informed consents were distributed to the fourth, fifth, and sixth grades of the school. Among those children/parents who signed the consent, a randomization was performed to select 60 who had availability to perform ~~two~~ separate physical and body composition assessments in our laboratory. Sixty-four students participated in the study, although only 41 finished the intervention. In summary, data for this analysis comes from a randomized sample from ~~two~~ primary schools in Spain, where PA assessment by questionnaires was one of the variables in those longitudinal studies.

Instruments

Initial measurements. Anthropometric measurements (waist and hip circumferences, sagittal abdominal diameter, and skinfolds), including height and body mass, were performed according to the International Society for the Advancement of Kinanthropometry (ISAK) standards for anthropometric assessment. Height was assessed with socks while shoes were taken off, using a stadiometer (SECA Leicester, Birmingham, UK). A Tanita UM-050 digital weighing scale (Tanita UK Ltd, Middlesex, UK) was used to assess body mass. Fat mass percent (FMP) was calculated using Slaughter's equation[14] from anthropometric measures. BMI was calculated using the classical equation and were categorized into ~~three~~ levels: normal-weight, overweight, and obesity, according Cole's ~~cut-off~~ points.[15]

Physical ~~activity questionnaire for children~~ (PAQ-C). PA was assessed using the PAQ-C.[5] The PAQ-C is a 9-item, 7-day PA recall designed for use with elementary and middle school children in a field-based setting. A 10th item not used in the calculation of the activity score asks children if they were sick or otherwise prevented from engaging in regular PA.

The PAQ-C was administered twice and children were asked to recall their participation in activities over the last 7 days to compute an activity score. Once a value from 1 to 5 for each of the ~~nine~~ items (items 1 to 9) used in the PA composite score is obtained, the mean of these ~~nine~~ items is taken, which results in the final PAQ-C activity summary score.

Cultural adaptation of the Spanish PAQ-C was performed following the basic steps of standardized questionnaires cultural adaptation process.[16] The research team members made the original Spanish translation. Subsequently, ~~two~~ bilingual researchers outside the group performed the reverse translation. The differences between the original version and the translations were reviewed and discussed by the research group and external researchers. Typically, the questionnaire was completed at school in a quiet room, and researchers were available to help children and confirm that all items were answered. The entire process lasted for approximately 10 to 15 minutes.

Accelerometry. The Actigraph GT3X monitor device (Actigraph, Pensacola, FL), was used to assess PA. The accelerometer is lightweight (27 g), compact ($3.8 \times 3.7 \times 1.8\,cm^3$) and has a rechargeable lithium polymer battery. It uses a solid-state triaxial accelerometer to collect motion data on ~~three~~ axes: vertical (Y), horizontal right-left (X), and horizontal front-back axis (Z). The GT3X measures accelerations in the range of 0.05 to 2 g, which is digitized by a 12-bit analog-to-digital converter at a rate of 30 Hz. Once digitized, the data are filtered using a band-limited frequency of 0.25 to 2.5 Hz. The Actigraph accelerometer has been shown to be a reliable and valid tool for the assessment of different types of physical activities.[17,18]

Researchers distributed ~~pre-initialized~~ accelerometers face-to-face at schools. Participants wore the accelerometers on the right side of the hip, secured with an adjustable elastic belt, underneath clothing, near to the center of gravity. Participants received a demonstration from a trained researcher on how to wear the accelerometer. They were asked to only remove the device when sleeping and engaging in water-based activities. In addition, children received a brochure about accelerometer use including the instructions. Accelerometers were set to register 1-second epoch cycles, and were programmed to start the record at midnight of the following day they received the monitor and to record activity for the following 7 days.

The version 6.11.1 of Actilife Software (Actigraph) was used to process the accelerometer data. Periods of ≥60 minutes of zero values, allowing for 2 minutes of ~~non-zero~~ interruptions, were defined as accelerometer "~~non-wear~~" time and were removed from the analyses. The first day of recording was not included in the analysis. Only participants with ≥4 complete days, including ~~one~~ weekend day, were included.[19] A day was considered valid if it contained ≥10 hours of wear time for weekdays and ≥8 hours for weekend days considering different sleep patterns over weekends.[20]

We selected the cut points from Evenson et al[21] to determine the time spent on different intensity levels of PA: ≤100 cpm for sedentary behavior, <2296 cpm for light, <4012 cpm for moderate, and ≥4012 cpm for vigorous PA.

A recording of more than 15,000 counts per minute was considered as a potential malfunction of the accelerometer and the value was excluded from the analyses, based on the recommendations from Esliger et al.[22]

Procedure

Following agreement to participate in the study, participants were assessed (initial measurements) and received an accelerometer and later (8 days) completed PAQ-C twice, with a gap of 6 hours between the ~~two~~ questionnaires. After the trial period, the material and questionnaires were collected by the researcher, and the data was stored in a spreadsheet using Microsoft Excel for further analysis.

Data Analysis

The characteristics of subjects were described with frequency distribution and mean, standard deviation (SD). The reliability (within-subject variability) was calculated by applying the PAQ-C during the same day twice and ~~intra-class~~ correlation coefficient (ICC) was used to confirm the reproducibility. In addition, agreement analysis was performed between first and second measurements of total score, also systematic and proportional bias were calculated by independent sample ~~t~~-test and Kendall's tau rank correlation. Individual item reliability was also carried out by the same procedure. The internal consistency of the questionnaire was analyzed using Cronbach's α coefficient. Removing every item to confirm or exclude redundancy of the individual items was also performed by Cronbach's α. In addition, we carried out an inter-item raw correlation coefficient to complete reliability analysis as suggested by Clark and Watson.[23]

The relationship between the PAQ-C and the accelerometer scores was performed using Spearman's rank correlation coefficients (rho). Agreement between the PAQ-C (total score values) and the accelerometer (MVPA minutes per day) was assessed using the Bland and Altman method;[24] after Z score transformation for PAQ-C and accelerometer values. The Bland and Altman plots give an indication of random error and bias. The first PAQ-C was selected always as reference, and the second one was used to carry out the ~~intra-day~~ and/or ~~intra-subject~~ reliability.

Table 1. **Characteristics of Study Participants By Sex**

	All (N = 83)	Boys (N = 46)	Girls (N = 37)
Age (years)	10.98 ± 1.17	11.09 ± 1.18	10.85 ± 1.16
Weight (kg)	42.06 ± 11.32	41.70 ± 10.58	42.51 ± 10.30
Height (cm)	145.85 ± 10.05	145.43 ± 10.11	146.38 ± 10.09
BMI (kg/m^2)	19.49 ± 3.51	19.44 ± 3.17	19.54 ± 3.94
Normal-weight (N = 49; 26♂, 23♀)	17.05 ± 1.75	17.15 ± 1.64	16.93 ± 1.90
Overweight (N = 28; 17♂, 11♀)	22.29 ± 1.43	21.84 ± 1.26	22.98 ± 1.45
Obesity (N = 6; 3♂, 3♀)	26.33 ± 1.88	25.70 ± 2.17	26.95 ± 1.73
Fat mass percent (%)	23.26 ± 7.63	21.62 ± 6.92	25.30 ± 8.06
PA total score (PAQ-C)	3.24 ± 0.64	3.22 ± 0.70	3.26 ± 0.56
MVPA (minutes/day)*	63.22 ± 14.40	63.43 ± 15.23	62.98 ± 13.57
Steps/day*	10,690.55 ± 1934.87	10,852.04 ± 2130.52	10,502.15 ± 1688.50

Values are presented as mean ± SD.
BMI, body mass index; PA, physical activity; MVPA, moderate-vigorous physical activity.
PA total score (PAQ-C) is the average of the ~~nine~~ items (items 1 to 9).
*N = 78; 42♂, 36♀.

Table 2. **Reliability Analyses of Physical Activity Questionnaire for Children (PAQ-C) by ~~Intra-Class~~ Correlation Coefficient (ICC; N = 83)**

	All	Boys (N = 46)	Girls (N = 37)	NW (N = 49)	OW + O (N = 34)
Total score	0.96	0.96	0.97	0.97	0.95
Item 1: activity checklist	0.96	0.94	0.98	0.97	0.94
Item 2: physical education	0.95	0.95	0.99	0.97	0.92
Item 3: recess	0.79	0.66	0.92	0.77	0.83
Item 4: lunch	0.87	0.87	0.89	0.84	0.90
Item 5: after school (14 to 18 hours)	0.82	0.80	0.86	0.79	0.84
Item 6: afternoon (18 to 22 hours)	0.77	0.69	0.88	0.71	0.87
Item 7: weekend	0.63	0.45	0.85	0.60	0.65
Item 8: intensity last week	0.90	0.91	0.89	0.89	0.92
Item 9: week summary	0.95	0.95	0.94	0.94	0.95

NW, normal-weight; OW + O, overweight plus obesity.
The reliability was calculated by ICC from ~~two~~ repeated measures by applying twice the PAQ-C during the same day. The first PAQ-C was selected always as reference.
$ICC = \frac{s_b^2}{(s_b^2 + s_w^2)}$

In respect to statistical power analysis, if we assume an 80% of power and α value of 0.001, it will permit us to detect coefficients of correlation as low as 0.3 with a sample size of 66 students, which is below our sample size, so we could confirm that our correlation analysis for validity will not be biased by the sample size (type II error). The analyses were done using SPSS 21.0 (SPSS Inc. Chicago, IL) and the level of significance was set at $p < .05$.

RESULTS

Characteristics of the participants are reported as mean and SDs in Table 1. No significant sex differences were found in the variables shown. Table 2 shows results for PAQ-C within-subject reliability. Low differences between the first and second questionnaire administration were observed (difference between first and second time scores = 0.0248 ± 0.2568). Also, the highest reliability values for the total score were obtained from girls and normal-weight children (ICC = 0.97 for both). Regarding individual item analyses, item 7, which informs about PA during the weekend, showed an ICC of 0.63, which was the lowest value. Moreover, compared with girls, the results reported by boys had an even lower ICC value (0.45 vs 0.85) in this particular item, so boy's answers biased the reliability in this item. Conversely, item 1 had the best reliability (ICC = 0.96) without differences between sexes.

The internal consistency coefficients of the questionnaire can be seen in Table 3. The PAQ-C questionnaire obtained a consistency of Cronbach's $\alpha = 0.76$. The internal consistency was reduced after removing every single item (Table 3), which may indicate that no redundant items are included in the questionnaire. Removing third and eighth items had the lowest values in the internal consistency, showing a high contribution to final test score, except for the group of overweight and obese children, who reported higher values in these items. The inter-item correlation coefficients were between 0.581 (items 1 and 9) and 0.177 (items 1 and 3), and the average inter-item correlation coefficient rho = 0.442.

Table 3. **The Internal Consistency of the Physical Activity Questionnaire for Children (PAQ-C) by Cronbach's α Coefficient (N = 83)**

	All	Boys (N = 46)	Girls (N = 37)	NW (N = 49)	OW + O (N = 34)
Total score	0.76	0.78	0.75	0.77	0.86
Item 1: activity checklist	0.58	0.55	0.65	0.55	0.58
Item 2: physical education	0.40	0.45	0.31	0.43	0.71
Item 3: recess	0.34	0.38	0.36	0.40	0.58
Item 4: lunch	0.44	0.60	0.16	0.46	0.48
Item 5: after school (14 to 18 hours)	0.52	0.45	0.65	0.51	0.46
Item 6: afternoon (18 to 22 hours)	0.59	0.53	0.69	0.51	0.67
Item 7: weekend	0.41	0.32	0.58	0.65	0.55
Item 8: intensity last week	0.38	0.53	0.13	0.25	0.72
Item 9: week summary	0.68	0.65	0.75	0.63	0.71

NW, normal-weight; OW + O, overweight plus obesity.
Internal consistency was calculated by Cronbach's coefficient α, which is based on the average correlation among the items in the questionnaire and the variance of the total score of PAQ-C. Cronbach's α coefficient by item indicates the α when that item is removed from the questionnaire. $\alpha = \frac{K}{K-1}\left[1 - \frac{\sum S_i^2}{S_T^2}\right]$

Table 4. **Spearman's Rank Correlations Coefficient (rho) for Physical Activity Questionnaire for Children (PAQ-C) and Total Physical Activity and Moderate-Vigorous Physical Activity (MVPA) Assessed by Actigraph (N = 78)**

	Total PA	MVPA/Day	Steps/Day
Total score	0.278*	0.248*	0.228*
Item 1: activity checklist	0.079	0.040	−0.030
Item 2: physical education	0.047	0.041	0.005
Item 3: recess	0.142	0.188	0.150
Item 4: lunch	0.074	0.001	0.022
Item 5: after school (14 to 18 hours)	0.145	0.144	0.161
Item 6: afternoon (18 to 22 hours)	0.282*	0.268*	0.203
Item 7: weekend	0.123	0.079	0.084
Item 8: intensity last week	0.243*	0.208	0.255*
Item 9: week summary	0.295**	0.304**	0.311**

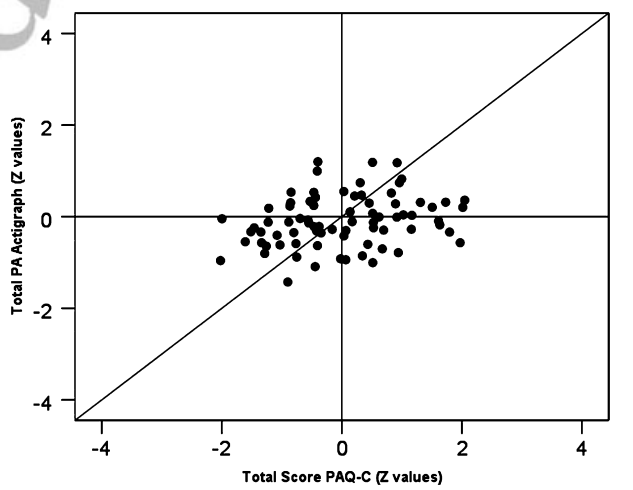
PA, physical activity.
*p < .05; **p < .01.

Spearman rank correlations between the PAQ-C and the accelerometer for the total PA were significant, although they remain weak (rho = 0.23-0.28, all p < .05). Items 6, 8, and 9 showed significant correlations with the total activity, MVPA and steps measured by accelerometry. The highest correlation was observed for item 9 (rho = 0.31, p < .01) (Table 4).

The accuracy analysis performed with the concordance coefficient correlation reported a low accuracy compared with accelerometry (r = .192, p = .092), which was graphically confirmed by the scatter plot between PAQ-C and accelerometry where regression line Z values of both instruments ($R^2 = .07$, p < .05) were compared with the identity line (Figure 1).

In Figure 2, the Bland and Altman plot for Z values of the total PA estimated by PAQ-C and accelerometry showed significant differences between the PAQ-C and the accelerometer. In addition, differences between the PAQ-C and the accelerometer scores increased as Z values, so higher Z mean, higher Z difference (Kendall's tau = 0.189, p < .05).

Figure 1. **Scatter Plot Between Z Values of Physical Activity Questionnaire for Children (PAQ-C) and Total PA by Actigraph**

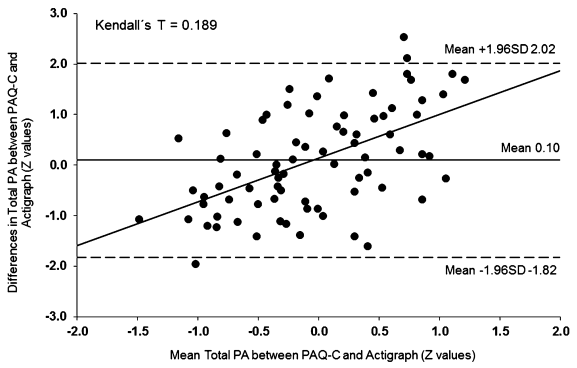


Note: Z value = (individual score − mean score)/SD.

DISCUSSION

This study evaluated the reliability and validity of the Spanish version of the PAQ-C questionnaire to assess PA in schoolchildren. One of the strengths of this study was to evaluate the reliability and validity of the PAQ-C using a triaxial accelerometer as criterion (setup: 1-second epoch). Also, it is the first time that this has been conducted with Spanish children. The results show high test-retest reliability (ICC = 0.96) for the Spanish adaptation of the PAQ-C. The reliability of the PAQ-C was good in other studies conducted in other countries[5,9,25] and with children of different races.[11] An estimate of Cronbach's coefficient α > 0.70 is usually considered indicative of a reliable questionnaire.[26] Our results confirm the satisfactory internal consistency of the questionnaire (α = 0.76), which is in accordance with other studies reporting similar results, so Crocker et al[5] reported 0.79 to 0.89

Figure 2. **Bland-Altman Plot With Differences in Mean Total PA for the Physical Activity Questionnaire for Children (PAQ-C) and the Actigraph**



Dashed lines are agreement limits; thin solid line represents mean difference between Z-scores of physical activity measured by PAQ-C and accelerometers; thick solid line indicates trend from Kendall's tau correlation., Z value = (individual score − mean score)/SD.

for ~~two~~ PAQ-C measurements; Janz et al[25] reported 0.72 to 0.76 after ~~three~~ repeated measures; and Moore et al[11] found an α of 0.75 for European American children.

However, the individual item analysis revealed that reliability was not uniform along the questionnaire, so the seventh item revealed a low ICC, which appears to be sex-dependent as boys had significantly lower consistency than girls. This item asks for PA during the weekend, and it should be more variable and irregular than weekdays (usually children have a schedule during weekdays), so it could be more difficult for children to reproduce the same answer. Regarding the difference between boys and girls being more difficult to interpret, we can speculate that maturity and behavior differences could be plausible causes, however this hypothesis would need to be investigated (ie, girls may have more consistent activities during the weekends than boys or the representativeness weekend activities are more important for girls than for boys, which may help girls to give more consistent answers). In addition, it was important to observe that internal consistency was lower than 0.70 after removing items (Table 3), which may confirm all items were important for the internal consistency of final score (no redundancy). These results of reliability were confirmed by the inter-item coefficient correlation of 0.442, which is in the range of values suggested in the literature.[23]

The PAQ-C questionnaire was compared with an objective measure of PA to determine the concurrent validity of the PAQ-C using triaxial accelerometry. The Spearman correlation coefficients between the PAQ-C and accelerometer scores ranged from 0.047 to 0.295 (rho 0.28 for total PA score; $p < .05$), indicating a low correlation between both instruments. These correlations for total activity were slightly lower than those obtained in the study reported by Janz et al.[25] This result does not concur with a validation study of the Physical Activity Questionnaire for Adolescents (PAQ-A) conducted on Spanish adolescents,[27] which showed reasonable validity of the PAQ-A for this age range (rho = 0.39; $p < .001$). These data confirm a line of evidence that suggests PA questionnaires for adolescents correlated better with accelerometer scores than PA questionnaires for children.[6] Nonetheless, it could be argued that a small sample size did not permit us to detect high correlation coefficients. However, our post hoc statistical power analysis informed that our results must not be influenced by the sample size. Therefore, our sample of 83 participants permitted us to detect correlation coefficients lower than 0.3.

Accelerometers have been suggested as one of the best criterion measures for validation of self-report instruments of PA.[28,29] However, this practice has been criticized because of the fact that accelerometers and self-report instruments measure different things[30] and both instruments produce considerably different results in PA intensity in this population.[31] Accelerometry measures body movement, while questionnaires often ask respondents to rate the activities related to the effort or frequency. Moreover, the activities reported in the first item of the PAQ-C, skateboarding and cycling, are impossible to capture with the accelerometer because these instruments only capture locomotive activities where center of gravity has oscillation. Further, PAQ-C asks for frequency spent in physical activities, which is a subjective rating of exercise load and different measurement unit compared to accelerometer. We tried overcoming this limitation to compare both instruments with Z value transformation, nevertheless, the results were similar, and significant differences and low correlations were confirmed; therefore, the observed agreement between the PAQ-C (a measure of self-report) and accelerometer (an objective measure) should be interpreted in this light.

The output from accelerometers is a dimensionless unit commonly referred to as accelerometer counts. Researchers have attempted to calibrate these counts with energy expenditure in order to get a biological meaning to the output.[32] This has resulted in the publication of count thresholds relating to various categories of energy expenditure that allow researchers to summarize time spent in a given intensity of activity.[33] The availability of multiple cut points or equations has led to much confusion in the accelerometer literature.[34] We used the Evenson et al[21] cut points, recommended in Trost et al[35] comparative study to estimate time spent in sedentary, light-, moderate-, and vigorous-intensity activity in children and adolescents. Other cut points would have yielded different results. Nonetheless, the associations and differences with total PA will continue to be

the same because this variable must not be highly dependent on cut-off values.

The Bland-Altman plots showed that PAQ-C gave higher values of total PA than the accelerometer. Participants are likely to overestimate the frequency of activities and this is not reflected in real movement data. In other words, greater the time spent in a certain level of PA assessed by the accelerometer, the greater the difference between both methods.

Finally, another source of error must be related with the adiposity level. In adolescents, an overestimation of PA has been reported when assessed by self-reported tools.[36] In our sample, an interaction between reliability and BMI groups was not plausible because similar coefficients of correlations were found among normal, overweight, and obese children. Also, non-significant correlations were observed between BMI or FMP and differences in total PA for the PAQ-C and accelerometry. However, it could be speculated that such results could be influenced by a low rate of overweight and obese children, but in this study the prevalence of overweight and obesity (40.9%, Table 1) was similar to the Spanish prevalence (34.9%).[37] In summary, we can conclude that the overestimation of PA founded in overweight and obese adolescents was not confirmed in children.

Limitations

We could speculate the low correlations observed in this study might be because of sample size as well as the level of maturity associated with the age of the children. However, the sample size was similar to the original validation studies[5,9] and the statistical power analysis informs that our sample size is enough to detect even stronger correlations as proved in our post hoc analysis. Other limitations could explain these low associations, so there must be inherent subjectivity when individuals are asked to respond to questions about their behavior. Some issues as recalling errors, deliberate misrepresentations, social desirability, and other biases have been pointed out to be particularly important when dealing with children.[38] All these previous limitations may be hard to overcome with the current protocol of the questionnaire, which assume a self-report procedure. So, only an enhanced protocol, which includes additional control items or provides more help to children to fill the PAQ-C (ie, by showing actual references such as PA-related videos of children), would improve the validity of the questionnaire. However, we could not implement any of these strategies as we followed the published procedure thoroughly.

CONCLUSIONS

In summary, this was the first study that analyzes validity and reliability of PAQ-C in Spanish children, which may help to understand the meaning and applicability of the questionnaire. The results suggest that PAQ-C had a high reliability but questionable validity for assessing PA in our sample of Spanish children. These findings would suggest that the PAQ-C requires additional development to be a useful measure of PA in Spanish children. Therefore, PA measurement in children should not be limited to self-report measures solely and whenever possible an ACL or other capture-movement device should be used. This approach will allow us to obtain a better interpretation of the actual results of PA in which children are involved. In view of the known benefits of PA in this population,[39-41] there is a need to develop new self-report measures or validate other existing PA questionnaires.

IMPLICATIONS FOR SCHOOL HEALTH

Quantifying PA will be helpful to focus school and community interventions on those groups with unhealthy lifestyles. However, the assessment of PA will be time-consuming and impractical in several settings.

This manuscript presents novel data from a reliability and validity study of the PAQ-C in children. The development of an efficient, effective, low-cost, and accurate method of measuring habitual PA in school children is an important and useful part of physical education and educational research. However, school personnel must be aware of the limitations and advantages of different measurement instruments when assessing PA in students.

We suggest, as long as the budget is possible and in accordance with other studies,[42] the use of both instruments when assessing PA levels, especially in the initial assessment and the design of efficient and effective intervention programs to enhance PA levels in children. Using a self-reporting tool, information about the type and schedule of PA is obtained; in addition, movement sensors such as ACLs provide us the actual quantity of PA. Obtaining both types of information will permit a better interpretation of PA effectiveness in order to diagnose and prescribe solutions for inactive and sedentary behavior. For example, children reporting many hours a week playing soccer (PAQ-C assessment) and low levels of MVPA (ACLs assessment) should be advised to improve participation during the games and workouts or engage more PA practice. Conversely, children reporting high levels of MVPA and sedentary behavior (ACL assessment) and sport training out-of-school and not PA during recess and physical education (PAQ-C assessment) should be empowered to modify this deleterious habit during school time.

Human Subjects Approval Statement

The research protocol was reviewed and approved by the Ethics Committees of the Sports Medicine School, at Faculty of Medicine (Málaga, Spain) and the Spanish Ministry of Economy and Competitiveness, which are recorded as EMEFIFE-3-2011 and DEP2011-30565, respectively. The study was developed following the ethical guidelines of the Declaration of Helsinki-Seoul, last modified in 2008.

REFERENCES

1. Nishida C, Uauy R, Kumanyika S, Shetty P. The joint WHO/FAO expert consultation on diet, nutrition and the prevention of chronic diseases: process, product and policy implications. *Public Health Nutr*. 2004;7(1A):245-250.
2. Cordova A, Villa G, Sureda A, Rodriguez-Marroyo JA, Sanchez-Colladoe MP. Physical activity and cardiovascular risk factors in Spanish children aged 11-13 years. *Rev Esp Cardiol*. 2012;65(7):620-626.
3. Andersen LB, Riddoch C, Kriemler S, Hills A. Physical activity and cardiovascular risk factors in children. *Br J Sport Med*. 2011;45(11):871-876.
4. Welk GJ, Corbin CB, Dale D. Measurement issues in the assessment of physical activity in children. *Res Q Exerc Sport*. 2000;71(2 suppl):S59-S73.
5. Crocker PRE, Bailey DA, Faulkner RA, Kowalski KC, McGrath R. Measuring general levels of physical activity: preliminary evidence for the ~~physical activity questionnaire~~ for older children. *Med Sci Sports Exerc*. 1997;29(10):1344-1349.
6. Chinapaw MJM, Mokkink LB, van Poppel MNM, van Mechelen W, Terwee CB. Physical ~~activity questionnaires~~ for youth. A systematic review of measurement properties. *Sports Med*. 2010;40(7):539-563.
7. Chen SR, Lee YJ, Chiu HW, Jeng C. Impact of physical activity on heart rate variability in children with type 1 diabetes. *Child Nerv Syst*. 2008;24(6):741-747.
8. Biddle SJH, Asare M. Physical activity and mental health in children and adolescents: a review of reviews. *Br J Sport Med*. 2011;45(11):886-895.
9. Kowalski KC, Crocker PRE, Faulkner RA. Validation of the ~~physical activity questionnaire~~ for older children. *Pediatr Exerc Sci*. 1997;9(2):174-186.
10. Voss C, Ogunleye AA, Sandercock GRH. Physical ~~a~~ctivity ~~q~~uestionnaire for ~~c~~hildren and ~~a~~dolescents: English norms and cut-off points. *Pediatr Int*. 2013;55(4):498-507.
11. Moore JB, Hanes JC, Barbeau P, Gutin B, Trevino RP, Yin ZN. Validation of the ~~physical activity questionnaire~~ for older children in children of different races. *Pediatr Exerc Sci*. 2007;19(1):6-19.
12. Guinhouya BC, Samouda H, de Beaufort C. Level of physical activity among children and adolescents in Europe: a review of physical activity assessed objectively by accelerometry. *Public Health*. 2013;127(4):301-311.
13. Loprinzi PD, Smit E, Cardinal BJ, Crespo C, Brodowicz G, Andersen R. Valid and invalid accelerometry data among children and adolescents: comparison across demographic, behavioral, and biological variables. *Am J Health Promot*. 2014;28(3):155-158.
14. Slaughter MH, Lohman TG, Boileau RA, et al. Skinfold equations for estimation of body fatness in children and youth. *Hum Biol*. 1988;60(5):709-723.
15. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ*. 2000;320(7244):1240-1243.
16. Geisinger KF. Cross-cultural normative assessment: translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychol Assess*. 1994;6(4):304-312.
17. Sasaki JE, John D, Freedson PS. Validation and comparison of ActiGraph activity monitors. *J Sci Med Sport*. 2011;14(5):411-416.
18. Santos-Lozano A, Santin-Medeiros F, Cardon G, et al. Actigraph GT3X: validation and determination of physical activity intensity cut points. *Int J Sports Med*. 2013;34(11):975-982.
19. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc*. 2008;40(1):181-188.
20. Yildirim M, Verloigne M, de Bourdeaudhuij I, et al. Study protocol of physical activity and sedentary behaviour measurement among schoolchildren by accelerometry - cross-sectional survey as part of the ENERGY-project. *BMC Public Health*. 2011;11:182.
21. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. *J Sports Sci*. 2008;26(14):1557-1565.
22. Esliger DW, Copeland JL, Barnes JD, Tremblay MS. Standardizing and optimizing the use of accelerometer data for free-living physical activity monitoring. *J Phys Act Health*. 2005;2(3):366-383.
23. Clark LA, Watson D. Constructing validity: basic issues in objective scale development. *Psychol Assess*. 1995;7:309-319.
24. Bland JM, Altman DG. Statistical-methods for assessing agreement between 2 methods of clinical measurement. *Lancet*. 1986;1(8476):307-310.
25. Janz KF, Lutuchy EM, Wenthe P, Levy SM. Measuring activity in children and adolescents using self-report: PAQ-C and PAQ-A. *Med Sci Sports Exerc*. 2008;40(4):767-772.
26. Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency. *J Pers Assess*. 2003;80(1):99-103.
27. Martinez-Gomez D, Martinez-de-Haro V, Pozo T, et al. Reliability and validity of the PAQ-A questionnaire to assess physical activity in Spanish adolescents. *Rev Esp Salud Publica*. 2009;83(3):427-439.
28. Ward DS, Evenson KR, Vaughn A, Rodgers AB, Troiano RP. Accelerometer use in physical activity: best practices and research recommendations. *Med Sci Sports Exerc*. 2005;37(11):S582-S588.
29. Vanhelst J, Hardy L, Gottrand F, Beghin L. Technical aspects and relevance of physical activity assessment in children and adolescents in free-living conditions. *Arch Pediatr*. 2012;19(11):1219-1225.
30. Ham SA, Reis JP, Strath SJ, Dubose KD, Ainsworth BE. Discrepancies between methods of identifying objectively determined physical activity. *Med Sci Sports Exerc*. 2007;39(1):52-58.
31. Kavanaugh K, Moore JB, Hibbett LJ, Kaczynski AT. Correlates of subjectively and objectively measured physical activity in young adolescents. *J Sport Health Sci*. 2014;4(3):222-227.
32. Freedson P, Pober D, Janz KF. Calibration of accelerometer output for children. *Med Sci Sports Exerc*. 2005;37(11):S523-S530.
33. Rowlands AV. Accelerometer assessment of physical activity in children: an update. *Pediatr Exerc Sci*. 2007;19(3):252-266.
34. Welk GJ, McClain J, Ainsworth BE. Protocols for evaluating equivalency of accelerometry-based activity monitors. *Med Sci Sports Exerc*. 2012;44:S39-S49.
35. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of accelerometer cut points for predicting activity intensity in youth. *Med Sci Sports Exerc*. 2011;43(7):1360-1368.
36. Elliott SA, Baxter KA, Davies PS, Truby H. Accuracy of self-reported physical activity levels in obese adolescents. *J Nutr Metab*. 2014;2014:808659.

37. Sanchez-Cruz JJ, Jimenez-Moleon JJ, Fernandez-Quesada F, Sanchez MJ. Prevalence of child and youth obesity in Spain in 2012. *Rev Esp Cardiol (Engl Ed)*. 2013;66(5): 371-376.
38. Sirard JR, Pate RR. Physical activity assessment in children and adolescents. *Sports Med*. 2001;31(6):439-454.
39. Ekelund U, Luan JA, Sherar LB, et al. Moderate to vigorous physical activity and sedentary time and cardiometabolic risk factors in children and adolescents. *JAMA*. 2012;307(7): 704-712.
40. Janssen I, LeBlanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *Int J Behav Nutr Phys Act*. 2010;7:40.
41. Sothern MS, Loftin M, Suskind RM, Udall JN, Blecker U. The health benefits of physical activity in children and adolescents: implications for chronic disease prevention. *Eur J Pediatr*. 1999;158(4):271-274.
42. O'Connor J, Ball EJ, Steinbeck KS, et al. Measuring physical activity in children: a comparison of four different methods. *Pediatr Exerc Sci*. 2003;15(2):202-215.

QUERIES TO BE ANSWERED BY AUTHOR

IMPORTANT NOTE: Please mark your corrections and answers to these queries directly onto the proof at the relevant place. DO NOT mark your corrections on this query sheet.

Queries from the Copyeditor:

AQ1. Please confirm that given names (red) and surnames/family names (green) have been identified correctly

AQ2. Please check and confirm the hierarchy of section headings.

AQ3. Table 4 was not cited in the text. An attempt has been made to insert the table into a relevant point in the text -please check that this is OK. If not, please provide clear guidance on where it should be cited in the text.

AQ4. Please spell out ACL.

AQ5. Refs. [5] and [38] were identical. Hence, we have deleted Ref. [38] and references have been renumbered. Please check and confirm.

AQ6. As per style, three author followed by et al. should be used if there are more than six authros in the reference list. Please check and confirm if there are more than six authors in all the et al. references.