Rev Bras Cineantropom Hum

original article

DOI: http://dx.doi.org/10.5007/1980-0037.2017v19n3p270

Reproducibility and concurrent validity of the Physical Activity Questionnaire for Adolescents (QAFA) aged 10-14 years

Reprodutibilidade e validade concorrente do Questionário de Atividade Física para Adolescentes (QAFA) de 10 a 14 anos de idade

Alcides Prazeres Filho[1]
Arthur Oliveira Barbosa[1]
Gerfeson Mendonça[1,2]
José Cazuza de Farias Júnior[1,2,3]

Abstract – The aim of the study was to evaluate the reproducibility and concurrent validity of the Physical Activity Questionnaire for Adolescents (QAFA). The reproducibility study (repeated evaluations with two-day interval between them) included 171 adolescents (average age 12.3 years, SD = 1.1), and the validity study (reference method: "*Actgraph GT3X*" accelerometer) included 341 adolescents (average age 11.9 years, SD = 1.0). All participants were students enrolled in public schools of Joao Pessoa (PB), Brazil. The QAFA reproducibility was measured using the intraclass correlation coefficient (ICC) and the Kappa PABAK Coefficient. Validity was evaluated by the Spearman correlation, Bland and Altman plot, sensitivity and specificity analysis. Reproducibility was satisfactory (ICC = 0.73; 95%CI: 0.63 - 0.79; Kappa PABAK = 0.58), and higher in physical activity frequency items (eight activities with ICC ≥ 0.70) than in duration items (seven activities with ICC ≥ 0.70). The correlation between QAFA and accelerometer was low (rho = 0.37; p < 0.001). Sensitivity was high (from 79.3% to 90.4%) and specificity was low (from 29.9% to 50.6%). The agreement between QAFA and accelerometer was satisfactory (Bland-Altman). QAFA showed satisfactory reproducibility and validity and can be used to evaluate physical activity in younger adolescents in face-to-face interviews.

Key words: Adolescence; Data accuracy; Motor activity.

Resumo – *O objetivo deste estudo foi avaliar a reprodutibilidade e a validade de concorrente do Questionário de Atividade Física para Adolescentes (QAFA). O estudo de reprodutibilidade (medidas repetidas com dois dias de intervalo entre réplicas de aplicação) incluiu 171 adolescentes (média de idade = 12,3 anos, DP = 1,1) e o de validade (critério de referência: acelerômetro "Actgraph GT3X") 341 adolescentes (média de idade = 11,9 anos, dp = 1,0). Todos eram estudantes do ensino fundamental II, da rede pública de ensino de João Pessoa (PB), Brasil. A reprodutibilidade do QAFA foi estimada por meio do coeficiente de correlação intraclasse (CCI) e do índice de concordância Kappa PABAK. A validade foi avaliada pela correlação de Spearman, diagrama de Bland-Altman, sensibilidade e especificidade. A reprodutibilidade do QAFA foi satisfatória (CCI = 0,73, IC95%: 0,63 - 0,79; Kapa PABAK = 0,58), sendo maior para medida de frequência de prática (oito atividades com CCI ≥ 0,70) comparado à de duração (sete atividades com CCI ≥ 0,70). A correlação entre a medida do QAFA e a do acelerômetro foi baixa (rho = 0,37; p < 0,001). A sensibilidade foi elevada (79,3% a 90,4%) e a especificidade baixa (29,9% a 50,6%). A concordância entre a medida do QAFA e a do acelerômetro foi satisfatória (Bland-Altman). O QAFA apresentou níveis satisfatórios de reprodutibilidade e validade concorrente, podendo ser aplicado na forma de entrevista face-a-face para mensurar atividade física em adolescentes mais jovens.*

Palavras-chave: *Adolescência; Atividade motora; Confiabilidade dos dados.*

1 Research and Study Group in Epidemiology of Physical Activity. João Pessoa, PB. Brazil.

2 Postgraduate Associated Program in Physical Education - UPE/UFPB, Recife, PE. Brazil

3 Department of Physical Education, Federal University of Paraíba, João Pessoa, PB. Brazil.

Received: 26 October 2016
Accepted: 07 April 2017

INTRODUCTION

Measuring physical activity in children and adolescents is a major challenge because they are involved in activities of different domains (leisure, school and commuting) and with marked variations in intensity, type, frequency and duration along the days of the week, producing several possibilities of practice accumulation during the day[1]. Several methods have been used to measure physical activity, being classified as subjective (questionnaires, reminders, diaries, logs) and objective (movement sensors, heart rate monitoring, direct behavior observation, doubly marked water)[2, 3].

Questionnaires are still the most widely used instruments to measure physical activity due to the possibility of adjusting to the study objective, measuring a large number of people at the same time, high applicability, low cost and ability to specify the physical activities that were performed[2]. Several physical activity questionnaires for adolescents have been produced and tested in the last 20 years[4-6]. Nevertheless, no improvement in their reproducibility and validity levels have been observed[7, 8].

Data from systematic reviews have shown that most questionnaires available to measure physical activity in adolescents do not allow estimating the level of physical activity according to recommendations[7,8]; were tested on non-representative samples[8]; did not use objective measures as a reference criterion in validation[8]; presented low methodological quality[7], but above all, did not have their psychometric properties evaluated in more than one population and in studies with good methodological quality[7].

It has been recommended that researchers instead of developing new physical activity questionnaires for adolescents, should re-evaluate those that have been most promising (instruments with acceptable reproducibility and validity levels)[7], using more robust reference criteria, with accelerometers and in representative samples of different age groups and different locations. It is also important to identify the factors that contribute to the lower reproducibility and validity levels of these instruments. Such procedures would limit the creation of new instruments only for groups or subgroups that do not yet have valid instruments.

One of the questionnaires considered promising for the measurement of physical activity in adolescents is the Physical Activity Questionnaire for Adolescents – QAFA[9]. This instrument was adapted from the Self-Administered Physical Activity Checklist[5] and tested in adolescents aged 14-19 years in João Pessoa (PB), and presented high levels of reproducibility (ICC = 0.88; 95%CI: 0.84-0.91) and moderate levels of validity (rho = 0.61, $p < 0.01$)[9]. QAFA consists of a list of physical activities that can be calibrated (most practiced activities and their respective nomenclatures) according to the age group and the context of the target population, measuring type, frequency and duration of leisure, school and commuting activities. This tool allows estimating the total time, by group or type of practice (daily or weekly) and classifying the level of physical activity according to physical activity recommendations for health.

However, in the QAFA validation, a 24-h recall was used as the reference criterion and adolescents aged 14-19 years. Although there is no consensus gold standard for the validation of physical activity questionnaires, the use of objective measures such as the accelerometer[10], has been recommended. It should also be considered that physical activity patterns (preferences and duration, and how physical activity accumulates) are different between the early and late periods of adolescence[11,12]. Finally, it is known that younger adolescents have greater difficulty in estimating the physical activity parameters that are normally measured (frequency and duration of activities), due to their involvement in different practices intermittently[13] and to the fact that their operational memory is not yet well developed. Thus, it is not known that the reproducibility and validity levels observed in older adolescents are equally satisfactory in younger adolescents (10-14 years). In this context, the aim of this study was to estimate the reproducibility level and concurrent validity of the Physical Activity Questionnaire for Adolescents (QAFA) aged 10-14 years.

METHODOLOGICAL PROCEDURES

The analysis of QAFA reproducibility and validity is part of the research project called Longitudinal Study on Sedentary Behavior, Physical Activity, Food and Health of Adolescents - LONCAAFS study, whose main aim is to analyze the interrelations between level of physical activity, sedentary behaviors, eating habits, quality of life and health in adolescents. This research project was approved by the Human Research Ethics Committee of the Federal University of Paraíba (Protocol No. 024/13).

Two studies were carried out to evaluate the psychometric properties of QAFA, one for reproducibility and the other for concurrent validity. In both, the target population was elementary school students from public schools of João Pessoa (PB).

Reproducibility study

To estimate the QAFA reproducibility, a study was carried out with a representative sample of elementary school students from the 6th and 8th grades. The minimum sample size determination considered intraclass correlation coefficient - ICC equal to or greater than 0.20 (reproducibility), two questionnaire applications, type I error of 5%, type II error of 20% (power of 80%) and an increase of 30% to compensate for losses and refusals, resulting in a sample of 95 adolescents.

The sample was selected by conglomerates in two stages: i) systematic selection of eight schools, proportionally distributed by type (state, municipal) and city region (north, south, east, west); ii) random selection of 15 classes proportionally distributed by grade (6th and 8th). The definition of the number of classes to be visited was based on the division of the number of adolescents who needed to be interviewed by the average number of students expected to be found per class. The number of schools to be

visited was defined in order to achieve the lowest agglomeration effect (fewer classes per school).

Data collection was performed from August to October 2013 by a team previously trained and submitted to a pilot study under the same conditions that would be found in the validation study of instruments.

The QAFA is composed of a list of 19 physical activities, with the possibility that the adolescents could add two more activities and was applied as a face-to-face interview, with a two-day interval between replications. The following aspects were measured: type (yes or no), frequency (days / week) and duration (minutes / day) of activities practiced in the last week prior to data collection for at least 10 minutes at each occasion. A moderate to vigorous physical activity score (minutes / week) was produced from the sum of multiplications of frequency by the respective durations of each activity. Adolescents with physical activity practice equal to or greater than 300 minutes per week were classified as physically active and the others as physically inactive[14].

To characterize the sample, the following data were collected: sex (male and female), age in complete years (categorized in 10-11 and 12-14 years of age); skin color (white and non-white), maternal schooling (incomplete elementary school, complete elementary school and complete high school) and economic class (Methodology of the Brazilian Association of Research Companies – ABEP[15], grouped into classes A / B, C and D / E). Body mass index (BMI) was determined as follows: body mass [kg] / height [m]^2 and classified according to World Health Organization - WHO criteria[16], specific for gender and age: low weight (<-2sd), normal weight (> -2dp and <+ 1dp), overweight (> + 1dp) and obesity (> + 2dp).

To describe the sociodemographic characteristics and BMI classification, frequency distribution was used. The ICC was used to estimate the reproducibility of continuous QAFA measurements (frequency, duration of practice by activities and total time of MVPA practice), adopting values equal to or greater than 0.70 as satisfactory[7]. In order to estimate the reproducibility of categorical measures (practices vs. does not practice each physical activity; physically active vs. physically inactive), the Kappa PABAK concordance index (Prevalence and Bias Adjusted Kappa) was used, classified as follows: up to 0.20: poor; 0.21 to 0.40: mild agreement; 0.41 to 0.60: moderate agreement; 0.61 to 0.80: substantial agreement and; > 0.81: almost perfect agreement[17].

Validity study

For the validity study, data collected from a representative sample of 6th graders enrolled in public elementary schools of João Pessoa (PB), corresponding to for the first year of collection (2014) of the LONCAAFS study, were used. The sample size was determined for a prevalence study: population equal to 9,520 6th grade students; prevalence of the outcome of 50%; 95% confidence interval; maximum acceptable error of four percentage points; design effect (deff) equal to two and increment of 40% for possible losses and refusals, resulting in 1,582 adolescents.

The measure of physical activity using accelerometer was used as a reference criterion to evaluate the QAFA concurrent validity. A subsample of adolescents was invited to use accelerometers (n = 1,039). Of the 28 schools selected to compose the LONCAAFS study sample, 17 schools (10 municipal and 7 state) were randomly selected and proportionally distributed by size (number of students enrolled in the 6th grade) and geographic region of the municipality (north, south, east and west). All students selected were invited to use the accelerometer.

Data collection was performed from February to June and from August to December 2014. In addition to the same variables evaluated in the reproducibility study, physical activity measured by accelerometry was included in this study. Adolescents were instructed to use the "Actigraph GT3X" accelerometer for seven consecutive days, fixed by elastic belt to the waist, removing only to sleep, to take bath and to carry out fight and / or aquatic activities. To reduce the amount of data, the ActiLife 6.12 software was used. Valid data were considered to have used for seven days at least eight hours a day. The presence of consecutive zeros in the data record for time longer than 60 minutes was considered as non-use. The epoch used was 15 seconds (reintegrated for one minute in the analyses)[13]. For the time of practice in MVPA, cutoff point was used (> 2,296 counts / minute)[13].

In order to analyze the relationship between MVPA time of QAFA and accelerometer, the Spearman correlation was used (data did not show adherence to normal distribution: $p < 0.05$ for the Kolmogorov-Smirnov test). The correlation coefficients were classified as weak magnitude correlation (< 0.40), moderate correlation (> 0.40 and < 0.50) and strong correlation (> 0.50)[18]. The absolute agreement between continuous time measurement of MVPA of QAFA and that of the accelerometer was evaluated using the Bland-Altman plot[19]. The difference between the measure of the questionnaire and that of the accelerometer (y-axis) and the mean between the two measures (x-axis) was plotted. The physical activity measures of QAFA and accelerometer were transformed into their respective logarithms, considering that this analysis requires that data do not present asymmetry and kurtosis at high levels[19]. They were considered satisfactory when the plotted values were within the limits of +1.96 sd [standard deviations] from the mean value of the difference of measurements evenly distributed on the 0 axis and without the presence of systematic error[19].

The QAFA sensitivity and specificity were calculated having as reference the accelerometer measurement. Sensitivity was defined as the proportion of adolescents classified as physically active (≥ 300 minutes / week) by QAFA who presented the same classification in the accelerometer measure, and specificity expressed the number of adolescents classified as physically inactive (< 300 minutes / week) by the QAFA and who presented similar practice in the measurement using the accelerometer.

RESULTS

Of the 250 adolescents selected to participate in the reproducibility study,

11 (< 10 or > 15 years of age) and 68 (not present in the second QAFA application) were excluded, resulting in 171 adolescents aged 10-14 years (mean of 12.3 years and SD = 1.1).

For the validity study, 1,039 adolescents were initially selected. Losses and refusals totaled 182 cases (15.9% of participants). Of the 857 adolescents who used accelerometer, 516 did not meet the data reduction criteria and the final sample was composed of 341 adolescents (mean age of 11.9 and SD = 1.0). In a later calculation, it was verified that this number of subjects is higher than the minimum necessary to identify correlations equal to or greater than 0.10, considering type I error of 5% and type II error of 20% (power of 80%), as statistically significant and sensitivity and specificity measures with values equal to or greater than 10%.

In both studies, the majority of adolescents were female, 10-11 years old, belonging to the economic class C / D / E, non-white, children of mothers with at least complete elementary school and about 30% were overweight (table 1).

For the reproducibility of physical activity type measurement, in two physical activities, Kappa PABAK values were classified as mild agreement, eight had moderate to substantial agreement, and six presented near perfect agreement. The ICC values ranged from 0.08 to 0.88 for the frequency measure and from 0.37 to 0.94 for the duration of the activities practiced. In eight activities, ICC values were above 0.70 for the frequency measure, while in seven activities, ICC exceeded 0.70 (table 2).

Table 1. Sociodemographic and BMI characteristics of adolescents for the reproducibility and validity study, João Pessoa (PB)

Variables	Reproducibility (n = 171)		Validity (n = 341)	
	n	%	n	%
Sex				
Male	69	40.4	169	49.6
Female	102	59.6	172	50.4
Age group (years)				
10-11	86	50.3	231	67.7
12-14	85	49.7	110	32.3
Economic class[a]				
A/B	51	31.9	116	39.2
C/D/E	109	68.1	180	60.8
Skin color				
Non white	129	24.6	277	81.2
White	42	75.4	64	18.8
Maternal schooling[b]				
Incomplete high school	56	40.8	99	35.6
Complete high school	30	19.2	74	26.6
Complete high school or more	71	40.0	105	37.8
BMI classification[c]				
Low weight	8	4.8	15	4.4
Normal weight	104	62.3	210	62.3
Overweight	36	21.5	67	19.9
Obesity	19	11.4	45	13.4

Unresponsive variables: Reproducibility study: a = 11; b = 14; c = 4; Validity study: a = 45; b = 63; c = 4

Table 2. Reproducibility of measurements of type, daily time, frequency and weekly duration of physical activity practice of adolescents, João Pessoa (PB), 2013

Physical activities	Number of practitioners	Practice? (yes or no)	Frequency (days / week)		Duration (Minutes / day)	
	n (%)	Kappa PABAK	ICC	95% CI	ICC	95% CI
1. Bicycle	67 (39.8)	0.79	0.80	0.73-0.85	0.63	0.50-0.73
2. Basketball	14 (8.1)	0.85	0.55	0.39-0.67	0.45	0.25-0.59
3. Walked as exercise	35 (20.5)	0.73	0.72	0.61-0.79	0.65	0.53-0.74
4. Running	31 (18.1)	0.37	0.08	-0.25-0.33	0.37	0.14-0.53
5. Soccer	63 (36.8)	0.55	0.74	0.64-0.81	0.66	0.53-0.75
6. Beach Soccer	7 (4.1)	0.93	0.78	0.71-0.84	0.84	0.79-0.88
7. Indoor soccer	67 (39.2)	0.60	0.65	0.52-0.74	0.78	0.70-0.84
8. Handball	26 (15.2)	0.78	0.60	0.45-0.71	0.60	0.46-0.71
9. Active commuting to School	101 (59.1)	0.79	0.86	0.80-0.89	0.42	0.21-0.58
10. Active commuting to other places	61 (35.7)	0.21	0.39	0.17-0.55	0.48	0.30-0.62
11. Dance	46 (27.0)	0.61	0.78	0.71-0.84	0.81	0.74-0.86
12. Fights	20 (11.7)	0.88	0.88	0.83-0.91	0.94	0.92-0.96
13. Weight lifting	23 (13.5)	0.83	0.71	0.65-0.81	0.73	0.64-0.80
14. Swimming	15 (8.9)	0.83	0.47	0.28-0.61	0.57	0.42-0.68
15. Games	87 (50.9)	0.49	0.50	0.25-0.67	0.53	0.28-0.69
16. Beach volleyball	10 (5.9)	0.89	0.19	-0.10-0.40	0.93	0.91-0.95
17. Volleyball	53 (31.0)	0.59	0.55	0.39-0.67	0.70	0.59-0.78

Two physical activities were excluded from the analysis because they presented low frequency in the second application: artistic gymnastics (n = 5) and gymnastics (n = 1); ICC: Intraclass correlation coefficient; 95% CI: 95% confidence interval. n: number of adolescents who answered yes to participate in each physical activity.

The reproducibility of the practice time of MVPA in the QAFA was ICC = 0.73 (95% CI: 0.63-0.79). The reproducibility of the categorized physical activity measure (physically active vs. physically inactive) was moderate (Kappa PABAK = 0.58), with no significant differences between subgroups (sex, age group and economic class) - Table 3.

The magnitude of the Spearman's correlation coefficient between the continuous measure (minutes / week) of the questionnaire and that of the accelerometer was classified as weak (rho = 0.37, p <0.001). The concordance between the logs of the QAFA and accelerometer measured by the Bland-Altman plot was satisfactory. The mean difference in time (minutes per day) of MVPA practice measured by QAFA and accelerometer was 280.3 (minutes / week), with agreement limits ranging from -561.2 (-1.96 * sd + mean) to 1,120.5 (+ 1.96 * sd + mean). Virtually all values of differences (98.5%) were within the limits of agreement of + 1.96 * dp, considered satisfactory. The presence of a systematic error in the QAFA measurement was not identified (figure 1).

The sensitivity of the QAFA measurement was high (equal to 85.4%), while the specificity was low (43%), being higher in females (specificity: 50.2%, 95% CI: 44.2-57, 0) compared to males (29.9%, 95% CI: 22.2 - 38.1) - Table 4.

Table 3. Reproducibility of the measurement of total time in moderate to vigorous intensity physical activity of adolescents, Joao Pessoa (PB), 2013

Variables		Time in MVVPA (minutes / week)		Classification of the level of physical activity (Physically active vs. physically inactive)
	n	ICC	95% CI	Kappa PABAK
All	171	0.73	0.63-0.79	0.58
Sex				
Male	69	0.75	0.59-0.84	0.48
Female	102	0.71	0.57-0.80	0.63
Age group (years)				
10-11	86	0.71	0.59-0.80	0.56
12-14	85	0.76	0.57-0.87	0.64
Economic class				
A/B	51	0.71	0.61-0.79	0.55
C/D/E	109	0.75	0.63-0.82	0.61

MVPA: Moderate to vigorous physical activity; ICC: Intraclass correlation coefficient

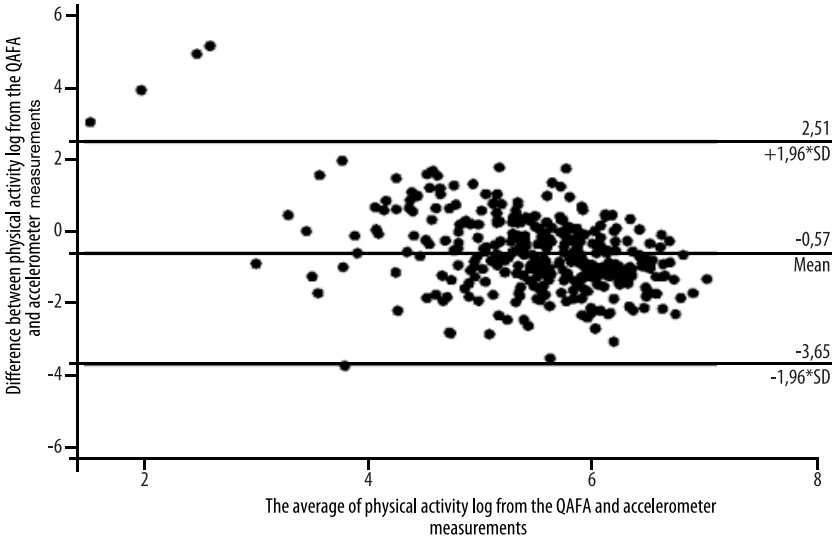

Figure 1. Bland-Altman scatter plot for absolute agreement between physical activity measure (MVPA time) from the questionnaire and accelerometer in adolescents, João Pessoa (PB, Brazil) 2014

Table 4. Spearman's correlation coefficient, sensitivity and specificity for the QAFA measurement using the accelerometer measurement as reference in adolescents, João Pessoa (PB), Brazil, 2014

Variables		MVPA time (minutes / week)	Categorical measure of physical activity (Physically active vs. physically inactive)			
	n	rho	Sensitivity	95% CI	Specificity	95% CI
All	341	0.37*	85.4	77.8-91.1	43.0	38.0-48.1
Sex						
Male	169	0.38*	87.2	78.8-93.2	29.9	22.2-38.1
Female	172	0.37*	79.3	60.3-92.0	50.6	44.2-57.0
Age group (years)						
10-11	231	0.36*	81.7	70.7-89.9	44.0	38.0-50.1
12-14	110	0.39*	90.4	79.0-96.8	40.7	31.6-50.4
Economic class						
A/B	116	0.37*	84.4	67.2-94.7	44.4	35.8-53.2
C/D/E	180	0.36*	84.4	74.4-91.7	40.4	33.6-47.6

MVPA: moderate to vigorous physical activity; *p < 0.001.

DISCUSSION

The aim of this study was to estimate the reproducibility and concurrent validity of QAFA in adolescents aged 10-14 years. In general, satisfactory reproducibility and concurrent validity levels were observed for the overall MVPA measurement of QAFA. However, the reproducibility levels were higher for the practice frequency measure compared to duration and for validity, correlation coefficients and specificity levels between physical activity measure produced by the questionnaire and accelerometer (reference criterion) were low. On the other hand, sensitivity levels and absolute agreement were high.

The QAFA reproducibility was 0.73 for continuous (minutes / week) and 0.58 for categorized measurement (physically inactive vs. physically active). These values are in agreement with those recommended in literature for physical activity measurements[7,8]. Helmerhorst et al.[20] analyzed questionnaires to measure physical activity in adolescents in the last seven days or week and verified that the reproducibility ranged from 0.58 to 0.92, and in only two instruments, it was higher than 0.70.

Chinapaw et al.[7] evaluated 31 physical activity for adolescents and verified that among the five questionnaires that were promising (composed of a list of moderate and vigorous physical activity activities, measured the frequency and duration of activities practiced in the last week: Oxford Physical Activity Questionnaire - OPAQ[21], Youth Physical Activity Questionnaire - YPAQ[22]), two were similar to QAFA. These instruments presented reproducibility levels (CCI > 0.85) slightly higher than those of the present study (CCI = 0.73). However, it is noteworthy that they were conducted in a convenience sample[22] from a single school[21] and with a more restricted age group (12 to 14 years)[21].

It was verified that the reproducibility levels for measurements type and frequency of physical activity practice were higher than duration and both were lower in non-systematized physical activities (e.g. active commuting to other sites). Telford et al.[23] evaluated the reproducibility of the Children's Leisure Activity Study Questionnaire (CLASS) in adolescents aged 10-12 years and observed higher ICC values for frequency (ICC = 0.36) compared to duration (ICC = 0.24) of practice. Lower reproducibility values for the measure of practice duration can be explained by the requirement of having a good working memory (ability to manipulate a limited set of information for a short time), which is still low and in development in this age group[24].

Another possible explanation is that adolescents, especially younger ones, have a pattern of physical activity characterized by involvement in multiple physical activities, with intermittent characteristics and marked variations throughout the day and between days of the week and the weekend[11]. Thus, memorizing the time of practice in these physical activities becomes an even more complex task for adolescents. In practice, the use of isolated information on the measure of the type of practice in the QAFA for sporadic activities, such as games, strolls and active commuting to other

places, should be analyzed with caution.

The magnitude of the correlation between QAFA and accelerometer measurements was weak (rho = 0.37, p < 0.001). In a systematic review of validation studies on physical activity questionnaires in adolescents performed by Helmershorsts et al.[20], the correlation coefficients between questionnaire and accelerometer measurements ranged from 0.06 to 0.50, and in 55% of questionnaires, the values of these coefficients were lower than those found in the QAFA. Questionnaires similar to QAFA (based on a list of physical activities practiced in the last week), when compared to accelerometer measures, presented correlation values close to the present study: Oxford Physical Activity Questionnaire - OPCW (r = 0.32)[21] and Youth Physical Activity Questionnaire - YPAQ (r = 0.42)[22].

The concordance between logs of the QAFA and accelerometer measurements was considered satisfactory (Bland-Altman plot). Although the mean difference between the two measurements was relatively high (280.3 minutes / week, 95% CI: -640.0 – 2,360), virtually all values were within the recommended agreement limits (+ 1.96 * standard deviation + mean), except for 1.5% (n = 5) of cases. No measurement bias was observed in the QAFA measure when compared to the accelerometer measurement (all points distributed along the 0 axis).

QAFA showed greater ability to adequately classify adolescents with sufficient levels of physical activity (high sensitivity levels: 79.3% to 90.4%) compared to those physically inactive (low specificity levels: 29.9% to 50.6 %). Some studies that included sensitivity and specificity analyses observed higher sensitivity (88.3%[25] and 80.0%[26]) and lower specificity levels (48.8%[25] and 57.0%[26]), even using accelerometer [26] as reference measure. Tavares et al.[25], in a validation study of physical activity indicators of the National School Health Survey questionnaire (PeNSE) found sensitivity values (77.9%) higher than those of specificity (69.1%), when compared to the measure of three 24-hour physical activity recalls. This can be explained by the fact that the reference criterion used was the 24-hour physical activity recall, which shows greater agreement with the questionnaire measure, compared to the accelerometer.

Information on sensitivity and specificity measures of physical activity questionnaires produced from the comparison with accelerometer measurements should be interpreted with caution, considering that it is not a gold standard and have basic differences with the measurements of questionnaires. However, low specificity values have been observed in other studies[25,26], which may be due to a possible overestimation of the questionnaire measure compared to the accelerometer measurement or vice-versa. In the case of the present study, the results indicated that adolescents with lower levels of physical activity (< 300min / week) in the accelerometer were not classified in the same way by the questionnaire.

Discrepancies between questionnaire and accelerometer measurements are expected and may have several explanations. First, they may come from the lack of homogeneity regarding the criteria required for validation of

accelerometer data (hours and days of accelerometer use, epochs and definition of non-use times, thresholds to define moderate to vigorous physical activity). Second, accelerometers did not measure some activities (aquatic activities, fights) and have limitations to measure others (weight lifting, cycling). Third, accelerometers can measure activities performed at short time intervals, for example, every minute, and those on a sporadic basis, while questionnaires, including that analyzed in the present study, consider only activities longer than or equal to ten minutes. Fourth, the questionnaire measured physical activities practiced in leisure, school and in commuting, while the accelerometer considered all activities practiced during the time of its use. It is not possible to rule out that some adolescents may not have unintentionally used the accelerometer during some physical activity practices. Finally, the questionnaire is susceptible to recall bias and presents high fragility in its practice duration measure, especially in young adolescents.

The strengths of this study include: representative sample of elementary school students from public schools of João Pessoa (PB) and; accelerometry as the reference criterion[20]. The use of concordance measures (sensitivity, specificity, Bland-Altman plot) and the evaluation of the reproducibility of measures of type, frequency and duration of physical activities are other strengths of this study.

Some limitations must be considered: the questionnaire applied as a face-to-face interview limits the generalization of the reproducibility and validity results for its application in another way. The non-inclusion of students from the private school system was another limitation, since adolescents from the private school system have higher level of education, income and parents' educational level, as well as higher levels of physical activity practice and greater participation in structured activities compared to those from the public school system[9]. These factors may interfere with the reproducibility and validity levels of the physical activity questionnaire. It is believed that if these adolescents were included in the study, the reproducibility and validity levels found in the present study would be higher.

It is important to stress that for some types of specific activities, the reproducibility of frequency level and practice duration were well below levels considered satisfactory (e.g., active commuting to other places). Thus, if the aim is to measure the overall level of physical activity, it is advised to use the QAFA. However, caution is required in the isolated use of some physical activities, since the reproducibility level obtained was low for some activities.

CONCLUSION

It was concluded that the overall QAFA measurement reached satisfactory reproducibility and concurrent validity levels, being similar to those observed in questionnaires classified as the most promising to measure physical activity in the last week in adolescents. Thus, the QAFA can be used as a face-to-face interview to measure physical activity in adolescents aged 10-14 years.

REFERENCES

1. Mindell JS, Coombs N, Stamatakis E. Measuring physical activity in children and adolescents for dietary surveys: practicalities, problems and pitfalls. Proc Nutr Soc 2014;73(02):218-25.
2. Reis RS, Petroski EL, Lopes AS. Medidas da atividade física: revisão de métodos. Rev Bras Cineantropom Desempenho Hum 2000;2(1):89-96.
3. Corder K, Ekelund U, Steele RM, Wareham NJ, Brage S. Assessment of physical activity in youth. J Appl Physiol 2008;105(3):977-87.
4. Guedes DP, Lopes CC, Guedes J, Stanganelli LC. Reprodutibilidade e validade do questionário Baecke para avaliação da atividade física habitual em adolescentes. Rev Port Ciênc Desporto 2006;6(3):265-74.
5. Sallis JF, Strikmiller PK, Harsha DW, Feldman HA, Ehlinger S, Stone EJ, et al. Validation of interviewer-and self-administered physical activity checklists for fifth grade students. Med Sci Sports Exerc 1996(28):840-51.
6. Monfort-Pañego M, Molina-García J, Miñana-Signes V, Bosch-Biviá A, Gómez-López A, Munguía-Izquierdo D. Development and psychometric evaluation of a health questionnaire on back care knowledge in daily life physical activities for adolescent students. Eur Spine J 2016;25(9):2803-8.
7. Chinapaw MJ, Mokkink LB, Poppel MN, Mechelen W, Terwee CB. Physical Activity Questionnaires for Youth. Sports Med 2010;40(7):539-63.
8. Farias Júnior JC, Lopes AS, Florindo AA, Hallal PC. Validade e reprodutibilidade dos instrumentos de medida da atividade física do tipo self-report em adolescentes: uma revisão sistemática. Cad Saude Publica 2010;26(9):1669-91.
9. Farias Júnior JC, Lopes AS, Mota J, Santos MP, Ribeiro JC, Hallal PC. Validade e reprodutibilidade de um questionário para medida de atividade física em adolescentes: uma adaptação do Self-Administered Physical Activity Checklist. Rev Bras Epidemiol 2012;15(1):198-210.
10. Aparicio-Ugarriza R, Mielgo-Ayuso J, Benito PJ, Pedrero-Chamizo R, Ara I, González-Gross M. Physical activity assessment in the general population; instrumental methods and new technologies. Nutr Hosp 2015;31(3):219-26.
11. Dumith SC, Gigante DP, Domingues MR, Kohl HW. Physical activity change during adolescence: a systematic review and a pooled analysis. Int J Epidemiol 2011;40(3):685-98.
12. Cooper AR, Goodman A, Page AS, Sherar LB, Esliger DW, Van Sluijs EM, et al. Objectively measured physical activity and sedentary time in youth: the International children's accelerometry database (ICAD). Int J Behav Nutr Phys Act 2015;12(1):1.
13. 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):1557-65.
14. WHO. Global recommendations on physical activity for health. World Health Organization, Geneva; 2010.
15. ABEP - Associação Brasileira de Empresas de Pesquisa. Critério de Classificação Econômica: Brasil. 2009; Disponível em: <http://www.abep.org/novo/Content.aspx?SectionID=84> [2014 jun 03].
16. WHO Multicentre Growth Reference Study Group. Child Growth Standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development. Geneva. 2006.
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33(1):159-74.
18. Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB. Delineando a pesquisa clínica-4: Artmed Editora; 2015.
19. Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. Lancet 1995;346(8982):1085-7.

20. Helmerhorst HHJ, Brage S, Warren J, Besson H, Ekelund U. A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. Int J Behav Nutr Phys Act 2012;9(1):1.
21. Lubans DR, Sylva K, Osborn Z. Convergent validity and test–retest reliability of the oxford physical activity questionnaire for secondary school students. Behav Change 2008;25(01):23-34.
22. Corder K, van Sluijs EM, Wright A, Whincup P, Wareham NJ, Ekelund U. Is it possible to assess free-living physical activity and energy expenditure in young people by self-report? Am J Clin Nutr 2009;89(3):862-70.
23. Telford A, Salmon J, Jolley D, Crawford D. Reliability and validity of physical activity questionnaires for children: The Children's Leisure Activities Study Survey (CLASS). Pediatr Exerc Sci 2004;16(1):64-78.
24. Gathercole SE. The development of memory. J Child Psychol Psychiatry 1998;39(1):3-27.
25. Tavares LF, Castro IRR, Cardoso LO, Levy RB, Claro RM, Oliveira AF. Validade de indicadores de atividade física e comportamento sedentário da Pesquisa Nacional de Saúde do Escolar entre adolescentes do Rio de Janeiro, Brasil. Cad Saude Publica 2014;30(9):1861-74.
26. Van Hoye A, Nicaise V, Sarrazin P. Self-reported and objective physical activity measurement by active youth. Sci Sports 2014;29(2):78-87.

CORRESPONDING AUTHOR

Alcides Prazeres Filho
Universidade Federal da Paraíba;
Centro de Ciência da Saúde,
Departamento de Educação Física;
Cidade Universitária, João Pessoa,
PB, Brasil.
CEP: 58051-900.
E-mail: alcidespf@hotmail.com