

Cuestionarios de actividad física para niños y adolescentes españoles: una revisión sistemática

Physical activity questionnaires for Spanish children and adolescents: a systematic review

R.I. Martínez-Lemos, C. Ayán Pérez, A. Sánchez Lastra, J.M. Cancela Carral, R. Valcarce Sánchez

RESUMEN

Los cuestionarios de valoración de la prevalencia de actividad física son herramientas muy empleadas por los profesionales de la salud. Es importante poner en su conocimiento la calidad y características de los mismos. Esta revisión sistemática tuvo como objetivo identificar y analizar las características y propiedades psicométricas de este tipo de cuestionarios cuando han sido aplicados a niños y adolescentes españoles. Se encontraron un total de 14 cuestionarios, obteniéndose valores de validez adecuados en dos de ellos. Solo dos cuestionarios obtuvieron una puntuación mayor que 5 sobre 10 en la evaluación de la calidad de sus propiedades de medida de acuerdo con la lista de comprobación *Qualitative Attributes and Measurement Properties of Physical Activity Questionnaires*. Los cuestionarios APALQ (9-18 años) y ENERGY (12-14 años), parecen ser, con sus limitaciones, los más adecuados de entre los disponibles para valorar la actividad física en niños y adolescentes españoles, contando ambos con limitaciones a considerar.

Palabras clave. Cuestionario. Actividad física. Niños. Adolescentes.

ABSTRACT

Questionnaires aimed at assessing the prevalence of physical activity are tools frequently used by health professionals. It is important to provide them with information about the characteristics and quality of these questionnaires. This systematic review aimed to identify and analyse the characteristics and psychometric properties of this type of questionnaires when administered to Spanish children and adolescents. A total of 14 questionnaires were found. Accurate validity values were obtained in two of them. Only two questionnaires obtained a score >5 in the "Qualitative Attributes and Measurement Properties of Physical Activity Questionnaires" checklist. Among the available questionnaires, the APALQ (9-18 years) and the ENERGY (12-14 years) seem to be the most accurate tools for assessing physical activity prevalence in Spanish children and adolescents, albeit with limitations that should be taken into account.

Keywords. Questionnaires. Physical activity. Children. Adolescents.

An. Sist. Sanit. Navar. 2016; 39 (3): 417-428

Departamento de Didácticas Especiales, Facultad de CC.EE. e de Deporte. Universidad de Vigo. Pontevedra. España.

Recepción: 09-05-2016

Aceptación provisional: 28-07-2016

Aceptación definitiva: 05-09-2016

Correspondencia

R. Iván Martínez Lemos
Facultad de CC.EE e de Deporte
Campus A Xunqueira, s/n
36005-Pontevedra
E-mail: ivanmartinez@uvigo.es

Sus autores declaran no tener ningún conflicto de intereses.

INTRODUCCIÓN

La promoción de la actividad física (AF) es una de las principales estrategias de salud a desarrollar entre la población escolar por sus beneficios sobre la salud en el presente, y por la gran probabilidad que existe de que los hábitos saludables adquiridos en la infancia se instauren y perpetúen en la edad adulta¹.

Para determinar la magnitud del efecto que dichas estrategias pueden llegar a alcanzar se hace necesario conocer el patrón de AF referido por la población diana, lo que supone que previamente han de ser identificadas herramientas adecuadas para tal fin².

A este respecto, los cuestionarios de valoración de AF han sido señalados como instrumentos útiles por su facilidad de administración, su bajo coste y su capacidad para extraer información de muestras numerosas en un corto intervalo de tiempo³. Sin embargo, carecen de la objetividad de otros métodos de valoración de la AF, como los acelerómetros, los podómetros, los pulsómetros o muy especialmente el agua doblemente marcada. Además pueden provocar sesgos de información debido a la capacidad de recuerdo limitada o alterada que las personas encuestadas pueden presentar⁴.

Por estos motivos, la elección de un cuestionario de AF como instrumento para analizar el patrón de AF debe realizarse atendiendo al comportamiento de dos propiedades psicométricas fundamentales, como son la fiabilidad y la validez⁵. Para ello, se considera de gran utilidad el estudio de las revisiones sistemáticas publicadas al respecto⁶.

En la literatura científica se vienen publicando revisiones sobre la utilidad del cuestionario como herramienta de valoración de la AF en población joven⁷⁻¹³. En ellas se analizan principalmente los estudios originales de creación y validación de un cuestionario, pero rara vez se incluyen evidencias sobre si ha sido traducido a otros idiomas. Esta ausencia de información sobre la adaptación cultural de los cuestionarios es una limitación impor-

tante a la hora de seleccionar una herramienta para la valoración de AF en niños y adolescentes españoles. Por este motivo, esta revisión sistemática tuvo como objetivo identificar y analizar las características y propiedades psicométricas de los cuestionarios de recuerdo de realización de AF aplicables en población de niños y adolescentes españoles.

MATERIAL Y MÉTODOS

Se llevó a cabo una revisión sistemática de estudios sobre cuestionarios de valoración de AF en población española de niños y adolescentes publicados en inglés o español sin acotación temporal. Se excluyeron los estudios que: a) no utilizasen una pregunta, serie de preguntas o cuestionario validado; b) no informasen de la validez y/o la fiabilidad de la pregunta, serie de preguntas o cuestionario empleado; c) utilizasen cuestionarios que evaluaran varios constructos además de la AF, salvo en el caso de que informasen sobre la validez y/o fiabilidad de las preguntas relacionadas específicamente con la práctica de la misma.

Se consultaron las siguientes bases de datos electrónicas: PubMed, Scopus, SportDiscus y Dialnet.

La estrategia de búsqueda fue diseñada para encontrar el mayor número de estudios que hayan tenido como objeto comprobar la validez y/o la fiabilidad de cuestionarios diseñados para valorar el patrón de AF en población escolar. Se realizó una primera búsqueda en tres bases de datos (PubMed, Scopus y SportDiscus), combinando las palabras clave *Physical Activity*, *Questionnaire Children*, *Adolescents* y *Spain*. Posteriormente se desarrolló una segunda búsqueda en las tres mismas bases de datos añadiendo a las palabras clave los términos *Validity* y/o *Reliability*. Finalmente, se repitió la estrategia de búsqueda en una cuarta base de datos (Dialnet), utilizando las palabras clave *Actividad Física*, *Cuestionario*, *Niños*, *Adolescentes*, *Validez* y *Fiabilidad*. En todos los casos se utilizó

el operador OR para combinar las palabras clave y se planteó un período de búsqueda sin acotación temporal.

Selección de estudios

Dos autores revisaron de manera independiente la información incluida en el título y/o el resumen de los registros tras la primera búsqueda (PubMed, Scopus y SportDiscus). Se realizó una primera selección en la que se incluyeron todas las investigaciones que tuvieron como objetivo principal analizar la validez y/o fiabilidad de una pregunta, serie de preguntas o cuestionario cuya respuesta ofreciese información sobre el patrón de AF realizado por población entre 5-18 años, identificándose la/s pregunta/s o el cuestionario empleado. No se incluyeron investigaciones que utilizaran cuestionarios que evaluaran varios constructos además de la actividad física. A este respecto, se localizó el texto completo de aquellos estudios que no presentasen en su título o resumen dicha información. En caso de desacuerdo consultaron a un tercer autor y se llegó a un consenso. Posteriormente, los dos mismos autores llevaron a cabo un segundo proceso de selección con el fin de localizar y revisar aquellos estudios, que aunque a juzgar por su título y/o resumen no tenían como objetivo informar sobre las propiedades psicométricas de cuestionarios o preguntas de cuantificación de AF, se intuyó que podrían aportar datos de relevancia al respecto. Finalmente, tres autores revisaron los textos completos resultantes del segundo proceso de selección para identificar el modo en que el patrón de AF fue valorado excluyéndose para un posterior análisis aquellas investigaciones que no utilizaran una pregunta, serie de preguntas o cuestionario validado para tal efecto, que no informasen de la validez y/o la fiabilidad de la pregunta, serie de preguntas o cuestionario empleado, que empleasen una pregunta, serie de preguntas o cuestionario previamente localizado en la primera selección o que utilizaran cuestionarios que evaluaran varios constructos

además de la actividad física. Las referencias bibliográficas de todos los estudios incluidos fueron también revisadas para identificar nuevas investigaciones que pudieran cumplir los criterios de inclusión anteriormente mencionados.

Extracción de datos

La información extraída de los estudios seleccionados se realizó mediante la lectura de los mismos por dos autores. Para cada cuestionario, pregunta o serie de preguntas se obtuvieron los siguientes datos: autor, país, lengua original de validación, proceso de adaptación cultural y capacidad para discriminar el grado de cumplimiento de las recomendaciones de consenso sobre nivel de AF para niños y adolescentes entre 5 y 17 años¹⁴. Para cada estudio se obtuvieron los siguientes datos: características de la muestra, variables evaluadas, grado de fiabilidad y/o validez informado así como el procedimiento empleado para su valoración incluido el intervalo de tiempo entre medidas y la descripción de la instrumentación empleada (podómetros, acelerómetros y pulsómetros).

Calidad metodológica

Se evaluó la calidad metodológica de los estudios incluidos mediante la lista de comprobación *Qualitative Attributes and Measurement Properties of Physical Activity Questionnaires (QAPAQ)*¹⁵. Este instrumento fue desarrollado para describir y evaluar de forma independiente los atributos cualitativos y las propiedades de medición de los cuestionarios de valoración de AF. Para describir los atributos cualitativos se emplearon 9 indicadores: 1. Constructo; 2. Población diana; 3. Escenario de práctica; 4. Procedimiento de administración; 5. Período de recuerdo; 6. Número de ítems; 7. Lista de actividades; 8. Dimensiones de práctica; 9. Unidad de medida. Para evaluar las propiedades de medida se emplearon 10 criterios que fueron calificados como ausentes o presentes. La puntuación

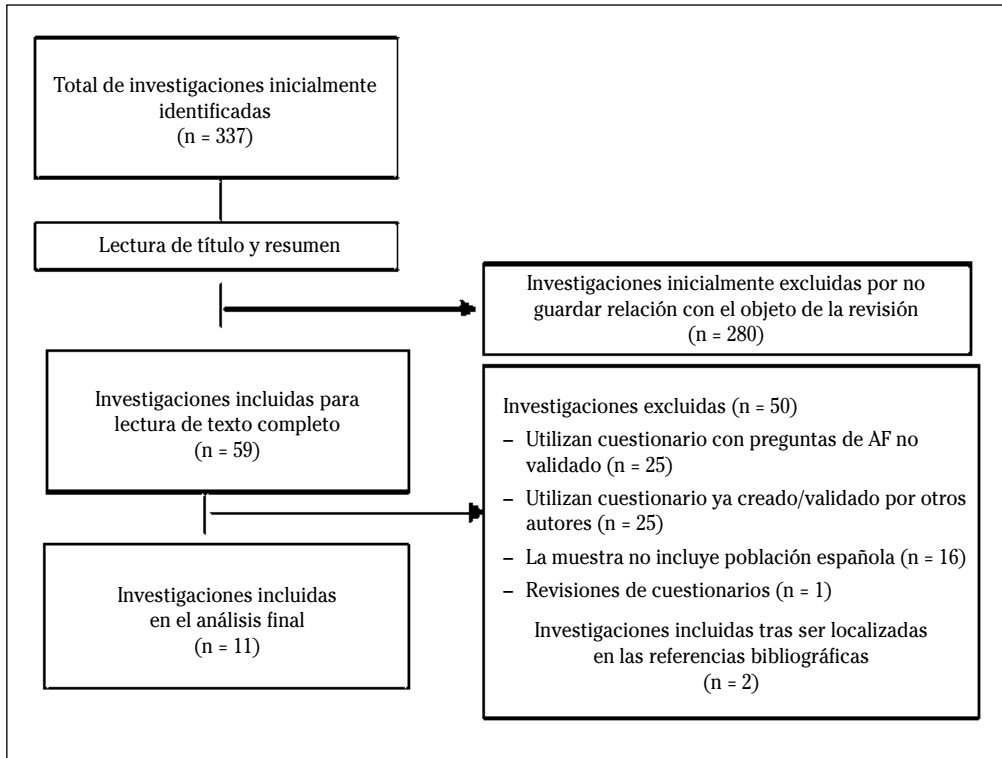


Figura 1. Procedimiento de selección de las investigaciones localizadas en el estudio.

máxima por estudio fue, por tanto, de 10 puntos. Los 10 criterios evaluados fueron: 1. Población de estudio; 2. Diseño del estudio; 3. Administración; 4. Fiabilidad-error de medida; 5. Fiabilidad-Test-retest; 6. Validez aparente; 7. Validez de contenido; 8. Efecto suelo-techo; 9. Validez de constructo y 10. Sensibilidad.

RESULTADOS

La búsqueda inicial permitió identificar 337 referencias y seleccionar 59 textos para su lectura completa, tras la cual se excluyeron 50 (Fig. 1). Simultáneamente la revisión de las referencias bibliográficas de los textos completos permitió localizar 2 estudios más. Así, finalmente se identificaron 11 estudios que informaron sobre la fiabilidad y/o la validez de 14 cuestionarios redacta-

dos en lengua castellana y empleados para cuantificar la AF realizada por niños y adolescentes (Tabla 1).

Algunos fueron diseñados originalmente para población de niños y adolescentes y otros son el resultado de adaptaciones de instrumentos originalmente creados para ser administrados en población adulta. Diez de los 14 cuestionarios analizados pertenecen al primer grupo, APALQ^{16,17}, 3-d BAD^{18,19}, FI-dPAQ^{20,21}, YACH^{22,23}, RPAR²⁴, Krece Plus^{25,26}, Fitnessgram^{25,27}, PAR^{25,28}, ENER-GY^{29,30} y ESCAF^{31,32}. Los cuatro restantes: IPAQ-A³³ adaptado de la versión larga del original IPAQ³⁴, PAQ-A³⁵ adaptado del original PAQ-C³⁶, en este caso de su versión corta, PDPAR-24³⁷ adaptado del PDPAR³⁸, y finalmente el PACE+²⁵ adaptado del original PACE³⁹, se encuadran en el segundo grupo. Solo dos cuestionarios (RPAR, Krece Plus)

Tabla 1. Lista de los cuestionarios para medir AF en castellano analizados

Acronimo	Definición y Referencias*	País ^a	Lengua ^b	Adaptación Cultural	Cumplimiento de recomendaciones ^c
APALQ	Assesment Physical Activity Levels Questionnaire ^{16,17}	Finlandia	Finlandés	Sí	No
3-d BAD	Bouchard Activity Diary ^{18,19}	Canadá	Francés	No	No
F1-dPAQ	Four by One Day Recall ^{20,21}	Reino Unido	Inglés	No	No
YACH	Yerterday Activity Checklist ^{22,23}	USA	Inglés	Sí	No
RPAR	Recess Physical Activity Recall ²⁴	España	Castellano	No	Sí
Krece Plus	Segunda pregunta del cuestionario Krece Plus ^{25,26}	España	Castellano	No	No
Fitnessgram	Primera pregunta del cuestionario Fitnessgram ^{25,27}	USA	Inglés	No	Sí
PAR	Physical Activity Rating ^{25,28}	USA	Inglés	No	No
ENERGY	Energy-Child Questionnaire ^{29,30}	Varios ^c	Varios ^c	No	No
ESCAF	Escala Semicuantitativa de AF ^{31,32}	Suecia	Sueco	No	No
IPAQ-A	International Physical Activity Questionnaire for Adolescents ^{33,34}	Varios ^d	Varios ^d	No	Sí
PAQ-A	Physical Activity Questionnaire for Adolescents ^{35,36}	Canadá	Francés	Sí	No
PDPAR-24	Previous Day Physical Activity Recall-24 ^{37,38}	USA	Inglés	Sí	Sí
PACE+	Patient-Centered Assessment and Counseling for Exercise Plus Nutrition ^{25,39}	USA	Inglés	No	Sí

* La primera referencia corresponde al estudio de validación al castellano del instrumento y la segunda al estudio de la versión original. ^aPaís original de validación; ^bLengua original de validación; ^cEl estudio de validez y fiabilidad se realizó simultáneamente en 6 países de Europa (Bélgica, Grecia, Hungría, Holanda, Noruega y España); ^dEl estudio de validez y fiabilidad se realizó simultáneamente en 12 países de Europa, América, Asia y Oceanía (Australia, Brasil, Reino Unido, Canada, Finlandia, Guatemala, Japón, Portugal, USA, Sudáfrica, Holanda y Suecia); ^eCapacidad para discriminar el grado de cumplimiento de las recomendaciones de consenso sobre el nivel de AF saludable para niños y adolescentes.

fueron creados originalmente para administrar a población escolar en lengua castellana. Todos los demás fueron desarrollados o adaptados a partir de una versión original en diferentes lenguas y solo 4 de ellos informaron de un método estándar de adaptación cultural⁴⁰. En la tabla 2 se muestra la descripción de los atributos cualitativos de los cuestionarios, siguiendo la estructura del QAPAQ. Como se puede observar en todos los casos se trató de autoinformes, salvo en el caso del F1-dPAQ y del ENERGY que fueron administrados mediante entrevista.

En la tabla 3 se describen las principales características de los estudios que analizaron la validez y fiabilidad de los cuestionarios analizados. Se informó de la fiabilidad test-retest de 7 cuestionarios, de los cuales YACH mostró los valores más altos (ICC=0,98) seguido del F1-d PAQ (ICC=0,84-0,93) y del RPAR (ICC=0,87-0,89). Se informó de la validez de constructo de 13 cuestionarios de los cuales el PDPAR-24 ($r=0,50-0,98$), F1-d PAQ ($r=0,67$) y YACH ($r=0,78$) mostraron los valores más altos frente a acelerómetro, podómetro y pulsómetro respectivamente.

Tabla 2. Descripción de los atributos cualitativos de los cuestionarios validados al castellano para medir AF en niños y adolescentes (n=14)

Cuestionario	Constructo	Población Diana y rango (años)	Escenario de práctica	Administración	Período de recuerdo	Formato			Interpretabilidad
						Items	Lista de Actividades	Dimensiones	Unidad de Medida
APALQ	AF habitual	9-18 (9)	Doméstico, Transporte, Escolar y Ocio	Autoinforme	7 días	5	No	T/F/D/I	Índice de AF (Escala 1-4)
3-d BAD	Gasto energético	12-16 (4)	Doméstico, Transporte, Ocio	Autoinforme	3 días	9	Sí	T/F/D/I/CS	Kcal/kg/15min
F1-dPAQ	Gasto energético	11-14 (3)	Doméstico, Transporte, Escolar y Ocio	Entrevista	4 días	24-28 ^b	Sí	F/D	Kcal/kg/día
IPAQ-A	Gasto energético	12-17 (5)	Doméstico, Transporte, Escolar y Ocio	Autoinforme	7 días	9	No	T/F/D/I/CS	MET-min/sem
PAQA	Gasto energético	13-18 (5)	Doméstico, Escolar, Ocio	Autoinforme	7 días	9	Sí	T/F	Puntuación (Escala 1-5)
PDPAR-24	Gasto energético	12-17 (5)	Doméstico, Transporte, Ocio	Autoinforme	1 día	69	Sí	T/F/D/I/CS	MET-min
YACH	Gasto energético	9-10 (1)	Doméstico, Ocio	Autoinforme	1 día	21	Sí	T/F	Frecuencia x MET
RPAR	Gasto energético	12-14 (2)	Doméstico, Transporte, Ocio	Autoinforme	1 día	NI	Sí	T/F/D/I/CS	MET-min
Krece Plus	Actividad deportiva extraescolar	4-14 (10)	Ocio	Autoinforme	7 días	1	No	T/D	Horas / sem
Fitnessgram	AF habitual	6-18 (12)	Transporte, Ocio	Autoinforme	7 días	1	No	T/F/D/I	Días /sem (≥ 60-min)
PACE+	AF habitual	11-18 (7)	Transporte, Ocio	Autoinforme	7 días	2	No	T/F/D/I	Días /sem (≥ 60-min)
PAR	AF habitual	11-18 (7)	NI	Autoinforme	Ninguno	1	No	NI	Índice de AF (Escala: 1-5)
ENERGY	Gasto energético	10-12 (2)	Doméstico, Transporte, Escolar y Ocio	Entrevista	7 días	30	No	T/F/D/I	Esc. de Intensidad (baja-media-alta)
ESCAF	AF habitual	8-13 (5)	Escolar, Ocio	Autoinforme	7 días	NI	No	T/F	Sesiones /sem (≥ 20-min)

^bCorrespondientes al cuestionario de jornada escolar y de fin de semana respectivamente. PA (procedimiento de administración); (AUT (autoadministrado); T/F/D/I/CS (Tipo/Frecuencia/Duración/Intensidad/Conducta Sedentaria); MET (Equivalente metabólico, unidad empleada para calcular el gasto de energía reflejado en el consumo de oxígeno de cualquier actividad física. MET (índice metabólico). NI (no informado).

Tabla 3. Análisis de validez y fiabilidad de los cuestionarios de medición de actividad física en población española de niños y adolescentes (n=14)

Cuestionario	Muestra de Estudio (tamaño, edad y género)	Variables evaluadas (Cuestionario & Instrumento)	Resultados	Intervalo (días)	Fiabilidad (test-retest)
			Validez		
APALQ ¹⁶	n=77 (13,6 ± 1,1 años) 7,7% chicas	IAF & AFMV (min/día)	Acelerómetro*; r=0,53 ^c (0,31 ^a -0,70 ^c)	7	—
	n=150 (13,6 ± 1,1 años) 52,0% chicas	IAF & AFM (pasos/día)	Acelerómetro*; r= 0,47 ^c (0,39 ^b 0,46 ^a)	7	ICC=0,76 (0,74-0,78)
3-d BAD ¹⁸	n=37 (12-16 años) 37,8% chicas	Total MET-min/día & Total counts/día	Acelerómetro*; rho=0,29	3	NI
F11-d PAQ ²⁰	n=37 (13,6 ± 1,1 años) 48,6% chicas	AFM & pasos/día	Podómetro+; r=0,67 ^b	7	ICC=0,93 ^a (Jornada Escolar) ICC= 0,84-0,91 ^a (Fin de semana)
		AFMV & (min > 50% FCR) & (min > 60% ICM) & (min > 140 ppm)	Pulsómetro≈ r=0,57 ^b ; r=0,59 ^b ; r=0,62 ^b	7	
IPAQ-A ³³	n=188 (15-17 años) NI % chicas	Total MET-min/día & AFVM (counts/min)	Acelerómetro*; rho=(0,20-0,27) ^b	4	NI
PAQ-A ³⁵	n=203 (13-17 años) 47,3% chicas	Ptos PAQ-A & Total counts/min Ptos PAQ-A & AFMV (≥ 1952 counts/min)	Acelerómetro*; rho= 0,39 ^c (0,34 ^b -0,17) Acelerómetro*; rho= 0,34 ^c (0,32 ^b -0,12 ^c)	7	—
	n=78 (12-17 años) 44,9% chicas		—	7	ICC= 0,71 (Puntuación Final)
PDPAR-24 ³⁷	n=79 (14-15 años) 39,2% chicas	AFLM	Acelerómetro**; r=0,50 ^a – 0,98 ^a	1	NI
YACH ²²	n=27; 10 años; NI % chicas	AFMV & (min > 120 ppm)	Pulsómetro¶; r= 0,78 ^c	1	—
	n=55; 10 años; NI % chicas		—	NI	ICC=0,98 ^a
RPAR ²⁴	n=49 (12-14 años) NI % chicas	Total AF & (counts)	Acelerómetro*; rho=0,43 ^c (0,35 ^c -0,32 ^b)	1	—
	n=32 (12-14 años) NI % chicas	Total AF & (pasos/día)	Podómetro+; r= 0,35 ^a ; kappa=0,21	1	—
		Total MET-min & Total count	Acelerómetro*; r= 0,40 ^a kappa=0,39	1	—
	n=113 (12-14 años) 47,8% chicas	Total MET-min & Total count	Acelerómetro*; r= 0,42 ^a ; kappa=0,16	1 hora	ICC= (0,87-0,89) ^a AF total
Krece Plus ²⁵	n=200 (13-17 años) 50,5% chicas	AF Total & (counts/min)	Acelerómetro*; rho=0,43 ^c	7	NI
Fitnessgram ²⁵	n=200 (13-17 años); 50,5% chicas	AF Total & (counts/min)	Acelerómetro*; rho=0,36 ^c	7	NI
PACE + ²⁵	n=200; 13-17 años; 50,5% chicas	AF Total & (counts/min)	Acelerómetro*; rho=0,43 ^c	7	NI
PAR ²⁵	n=200; 13-17 años; 50,5% chicas	AF Total & (counts/min)	Acelerómetro; rho=0,39 ^c	7	NI
ENERGY ²⁹	n=15; (11,3 ± 0,5 años) 47% chicas	Cuestionario & Entrevista	ICC= 0,46 (0,02-1,00)	7	—
	n=86; (11,3 ± 0,5 años) 44% chicas	Cuestionario & Entrevista	—	7	ICC= 0,65 (0,21-1,00)
ESCAF ³¹	n= 2255 (10-13 años)	AF Total	NI	28-56	τ = (0,44-0,60)

AFM (actividad física moderada); AFMV (actividad física moderada y vigorosa); AFLM (actividad física ligera y moderada); IAF (índice de actividad física); r=(Coeficiente de correlación de Pearson); rho= (Coeficiente de Correlación de Spearman); kappa= (Coeficiente kappa de Cohen); ICC=(índice de correlación intraclase); τ=(Tau-b de Kendall); NI (no informada); ^ap<0,05 ; ^bp<0,01; ^cp<0,001; * (ActiGraph GT1M); ** (ActiGraph GT3X); + (Yamax DW-450); = (Polar Accurex); ¶(Polar PE 4000).

Finalmente, en la tabla 4 se puede observar el resultado de aplicar el instrumento QAPAQ para evaluar las propiedades de medida de los instrumentos analizados. En ella

se observa que ninguno de los 14 cuestionarios analizados cumple el 100% de los criterios de calidad establecidos y que solo cinco mostraron un grado de cumplimiento $\geq 50\%$.

Tabla 4. Evaluación de las propiedades de medida de los cuestionarios de valoración de AF en población española de niños y adolescentes (n=14)

	General		Fiabilidad			Validez				Puntos sobre el Total	
	Población de Estudio ^a	Diseño del Estudio ^b	Adm ^c	Error de medida ^d	Test-retest ^e	Aparente ^f	Contenido ^g	Efecto suelo-techo ^h	Construido ⁱ		Sensibilidad ^j
APALQ ¹⁶	+	+	+	-	+	+	+	-	+	-	7/10
PAQ-A ³⁵	+	+	+	-	+	+	-	-	+	-	6/10
IPAQ-A ³³	+	-	+	-	-	+	+	-	+	-	5/10
PDPAR-24 ³⁷	+	+	+	-	-	+	-	-	+	-	5/10
F1-d PAQ ²⁰	+	-	+	-	+	-	+	-	+	-	5/10
RPAR ²⁴	+	-	+	-	+	-	-	-	+	-	4/10
YACH ²²	-	-	+	-	+	+	-	-	+	-	4/10
Krece Plus ²⁵	+	+	+	-	-	-	-	-	+	-	4/10
Fitnessgram ²⁵	+	+	+	-	-	-	-	-	+	-	4/10
PACE ⁺ ²⁵	+	+	+	-	-	-	-	-	+	-	4/10
PAR ²⁵	+	+	+	-	-	-	-	-	+	-	4/10
ENERGY ²⁹	+	-	+	-	+	-	-	-	+	-	4/10
3-d BAD ¹⁸	+	-	+	-	-	-	-	-	+	-	3/10
ESCAF ³¹	+	+	+	-	-	-	-	-	-	-	3/10

^aPoblación (descripción completa de la muestra; edad, género y nacionalidad); ^bDiseño del estudio (descripción de la versión del cuestionario usado, el tiempo entre mediciones y un tamaño de muestra adecuado $n > 50$ sujetos); ^c Administración (descripción del procedimiento de administración); ^dError de medida (Bland Altman, error estándar de medida, cambio mínimo detectable) ^eCoefficientes de fiabilidad (ICC, Kappa de Cohen). ^fValidez aparente o lógica (grado en que los ítems del cuestionario reflejan adecuadamente el constructo a medir y son entendibles); ^gValidez de contenido (grado en que todos los aspectos relevantes relacionados con el constructo a valorar han sido incluidos con suficiente detalles como contenido); ^hEfecto suelo-techo ($> 15\%$ de la muestra de estudio presenta los valores más altos o más bajos posibles, lo cual demostraría que el instrumento no es capaz de discriminar entre grupos y por lo tanto su fiabilidad es reducida y su sensibilidad limitada); ⁱValidez de Construido (comparación de los resultados del cuestionario con los de un instrumento validado que mide idéntico constructo); ^jSensibilidad (habilidad del instrumento para detectar cambios a lo largo del tiempo en el constructo medido). + (cumple el criterio); - (incumple el criterio)

DISCUSIÓN

Tras el análisis de los artículos seleccionados se constata que no todos los cuestionarios de recuerdo de realización de AF aplicables en población de niños y adolescentes españoles han evaluado conveniente sus propiedades de fiabilidad y validez. Así, del total de 14 cuestionarios analizados, 7 informaron acerca de su fiabilidad test-retest y 13 cuestionarios infor-

maron de su validez. Respecto al análisis de fiabilidad, todos los cuestionarios, salvo uno (ENERGY) mostraron valores superiores al estándar mínimo⁴¹ ($ICC=0,70$) y en al menos dos casos (YACH, F1-dPAQ) los valores mostrados podrían calificarse de alta fiabilidad⁴². Sin embargo, una limitación común a todos los estudios de fiabilidad es que ninguno incluyó la cuantificación del error de medida. Esta propiedad permite informar del error aleatorio o sistemático

en la puntuación del cuestionario por medio de parámetros como el límite del acuerdo, el tamaño medio del error y/o el denominado mínimo cambio detectable⁴³. Esta propiedad es especialmente importante en estudios que pretenden medir el efecto de una intervención⁴⁴.

En relación al análisis de validez se constató que ningún estudio informó sobre la validez de criterio de los cuestionarios. Esto es algo lógico, dado que en el caso de la medición de AF no existe un patrón oro^{9,45,46} y aunque para la medición de gasto energético total se ha sugerido que el agua doblemente marcada podría ser considerado como tal, se ha constatado que resulta una medida inapropiada para otras dimensiones de la AF, tales como la frecuencia, la intensidad, la duración o el tipo de AF^{15,44}. Así, del total de 14 cuestionarios analizados, 13 informaron de su validez de constructo, para la cual se acepta que el acelerómetro es el instrumento óptimo de comparación, frente a podómetros y pulsómetros⁴⁷ y aunque no hay consenso en torno a los valores de correlación que demostrarían una adecuada validez⁴⁵, diferentes autores han considerado un valor mínimo estándar^{48,49} ($r \geq 0,50$). Así, de los 13 cuestionarios que evaluaron validez de constructo, solo diez lo hicieron comparando resultados con acelerómetros y únicamente dos de ellos (PDPAR-24, APALQ) habrían obtenido resultados por encima del punto de corte.

En relación al tamaño de la muestra de los estudios para evaluar estas dos propiedades de medida, en general fueron reducidos, hasta el punto que cinco cuestionarios examinaron su validez y un sexto cuestionario validez y fiabilidad con un tamaño de muestra inferior a 50 sujetos, es decir por debajo del tamaño muestral mínimo aconsejable⁵⁰. El resultado de comparar los valores de las propiedades de medición de los cuestionarios en su versión en castellano con los obtenidos en los estudios originales es dispar. Respecto al análisis de fiabilidad; el valor obtenido con la versión española²² del YACH es superior al obtenido con la versión original²³ para muestras muy similares (ICC=0,98; $p < 0,05$; $n=55$; vs

ICC=60; $p < 0,05$; $n=66$). Lo mismo ocurre en el caso de la versión española del F1-dPAQ²⁰ que también mostró valores superiores a los obtenidos en el estudio original²¹, aunque en este caso las muestras no fuesen tan coincidentes (ICC=0,84-0,93; $p < 0,05$; $n=37$ vs ICC= 0,62; $p < 0,05$; $n=12$). En los otros tres casos no ha sido posible establecer comparaciones. Respecto al análisis de validez, el valor obtenido con la versión española²⁰ del caso del F1-dPAQ es ligeramente superior al obtenido con la versión original²¹ para muestras muy similares ($r=0,62$; $p < 0,01$; $n=37$; vs $r=0,61$; $p < 0,01$; $n=20$). En el caso del IPAQ-A, la versión validada al castellano³³ mostró unos valores similares a los obtenidos con la versión original³⁴ pero con una muestra mucho más pequeña ($r=0,23$; $p < 0,05$; $n=188$ vs $r=0,26$; $p < 0,05$; $n=1.696$). Para el resto de cuestionarios no fue posible establecer esta comparación.

En lo referente a los atributos cualitativos de los cuestionarios se constató también una gran disparidad. Así, un buen número de cuestionarios (APALQ, Krece Plus, Fitnessgram, PACE+, PAR, ESCAF) no presentaron una descripción lo suficientemente detallada del constructo que miden, mientras que solo cuatro (APALQ, F1-dPAQ, IPAQ-A, ENERGY), incluyeron los cuatro escenarios principales en los que se puede desarrollar la AF (doméstico, transporte, escolar y ocio) y de los que depende la precisión con la que los mismos pueden estimar la cantidad de AF realizada⁵¹. En relación a esto resulta inexplicable que 8 de los cuestionarios identificados (3-d BAD, PDPAR-24, YACH, RPAR, krece Plus, Fitnessgram, PACE+, PAR) no incluyesen el ámbito escolar como escenario específico a valorar. Respecto a la población diana, no todos los cuestionarios se dirigen a un grupo de edad concreto, ni diferencian entre niños y adolescentes, a pesar de que las dos principales fuentes de sesgo en el caso de estas herramientas de valoración son el nivel de desarrollo cognitivo y el grado de definición y comprensión de las variables elegidas para describir la AF⁵². Otras debilidades observadas en los cuestionarios estuvieron relacionadas con la ausencia de una definición clara que permita diferen-

ciar los diferentes tipos de intensidad de la AF realizada o la falta de información detallada sobre el tiempo estimado necesario para cumplimentarlo.

Entre los aspectos positivos a destacar se encuentra el hecho de que todos los cuestionarios tomaron como referencia temporal un periodo no superior a 7 días, aspecto que reduce la influencia del sesgo de recuerdo característico de este tipo de herramientas de valoración⁹. Entre los cuestionarios analizados solo 3 diferenciaron entre semana y fin de semana (F1-dPAQ, 3-d BAD, PAQ-A), algo que se considera de gran importancia para población en edad escolar^{15,53}. Igualmente otra característica positiva observada fue la posibilidad de que las puntuaciones de los cuestionarios pudieran ser expresadas en unidades estándar de gasto energético (MET-min; Kcal/kg/min) o convertidas a las mismas desde otras unidades de tiempo (duración, frecuencia) lo que permite identificar de manera rápida y directa, a partir del resultado obtenido en el cuestionario, si los encuestados cumplen con las recomendaciones publicadas al respecto para población de niños y adolescentes⁵⁴.

El resultado tras la evaluación de la calidad metodológica con un instrumento específico (QAPAQ) indica que APALQ y PAQ-A, cumplen respectivamente el 70 y el 60% de los requerimientos que se consideran óptimos en cuanto a las propiedades de medida de un cuestionario de valoración de AF. En consecuencia, proponemos la versión española del cuestionario APALQ¹⁶ como un instrumento adecuado para valorar la AF realizada por niños de entre 9 y 18 años. En menor medida, el cuestionario ENERGY puede resultar un instrumento aceptable pero restringido a edades comprendidas entre los 12 y 14 años. No obstante, se debe tener en cuenta que ambos instrumentos presentan importantes limitaciones y en todo caso serían apropiados para estudios transversales, no así para estudios longitudinales, habida cuenta que en ninguno de ellos se ha estudiado la sensibilidad para detectar cambios. En este sentido es necesario recordar que el tipo de diseño y el objetivo del estudio de valo-

ración de AF para el cual se emplean estos cuestionarios, condicionan en buena medida la elección de “patrón oro” de referencia y también la interpretación de la relevancia que se debe dar a los valores de las principales propiedades de medida.

BIBLIOGRAFÍA

1. KOHL HW, FULTON JE, CASPERSEN, CJ. Assessment of physical activity among children and adolescents: a review and synthesis. *Prev Med* 2000; 31: S54-S76.
2. WELK GJ (Ed.). *Physical activity assessment in health-related research*. Champaign IL: Human Kinetics 2002; 135 p.
3. SHEPHARD RJ. Limits to the measurement of habitual physical activity by questionnaires. *Br J Sports Med* 2003; 37: 197-206.
4. CORDER K, EKELUND U, STEELE RM, WAREHAM J, BRAGE S. ASSESSMENT OF PHYSICAL ACTIVITY IN YOUTH. *J App Physi* 2008; 105: 977-987.
5. AINSWORTH BE, CASPERSEN CJ, MATTHEWS CE, MÄSSE LC, BARANOWSKI D, ZHU, W. Recommendations to improve the accuracy of estimates of physical activity derived from self report. *J Phys Act Healt* 2012; 9: S76.
6. BETTANY-SALTIKOV J. Learning how to undertake a systematic review: part 1. *Nurs Stand* 2010; 24: 47-55.
7. TROST, SG. State of the art reviews: measurement of physical activity in children and adolescents. *Am J Lifestyle Med* 2007; 1: 299-314.
8. TESSIER S, VUILLEMIN A, BRIANÇON S. Review of physical activity questionnaires validated for children and adolescents. *Sci Sports* 2008; 23: 118-125.
9. CHINAPAW MJ, MOKKINK LB, VAN POPPEL MN, VAN MECHELEN W, TERWEE. Physical Activity Questionnaires for Youth: a systematic review of measurement properties. *Sports Medicine* 2010; 40: 539-563.
10. FARIAS JÚNIO JCD, LOPES ADS, FLORINDO AA, HALLAL, PC. Validity and reliability of self-report instruments for measuring physical activity in adolescents: a systematic review. *Cad de Saude Publica* 2010; 26: 1669-1691.
11. HELMERHORST HJ, 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: 103-157.
12. SLEDDENS EFC, KREMERS SPJ, HUGHES SO, CROSS MB, THIJIS C, DE VRIES NK, et al., Physical activity parenting: a systematic review of ques-

- tionnaires and their associations with child activity levels. *Obes Rev* 2012; 13: 1015-1033.
13. CANCELA JM, AYÁN C, CASTRO A. An evaluation of questionnaires assessing physical activity levels in youth populations. *J Child Health Care* 2013; 17: 274-279.
 14. WHO. Global recommendations on physical activity for health. Geneva: World Health Organization (WHO) 2010; 58 p.
 15. TERWEE CB, MOKKINK LB, VAN POPPEL MN, CHINAPAW JM, VAN MECHELEN W, DE VET HCW. Qualitative attributes and measurement properties of physical activity questionnaires. *Sports Medicine* 2010; 40: 525-537.
 16. ZARAGOZA J, GENERELO E, AZNAR S, ABARCA A, JULIÁN JA, MOTA J. Validation of a short physical activity recall questionnaire completed by Spanish adolescents. *Eur J Sport Sci* 2012; 12: 1-9.
 17. TELAMA R, VIKARI J, VALIMAKI I. Atherosclerosis precursors in Finnish children and adolescents: Leisure-time physical activity. *Acta Paediatr Scand* 1985; 318: 169-180.
 18. MARTINEZ-GOMEZ D, WARNBERG J, WELK GJ, SJÖSTROM M, VEIGA O, MARCOS A. Validity of the Bouchard activity diary in Spanish adolescents. *Public Health Nutr* 2010; 13: 261-268.
 19. BOUCHARD C, TREMBLAY A, LEBLANC C, LORTIE G, SAVARD R, THÉRIAULT G. A method to assess energy expenditure in children and adults. *Am J Clin Nutr* 1983; 37: 461-467.
 20. SARRADELL JJS, LANASPA EG, CASTERAD JZ, JULIAN JA. Validez de criterio y confiabilidad del "Four by One Day Physical Activity Questionnaire" en población adolescente española / "Validity and Reliability Criteria for the Four by One-Day Physical Activity Questionnaire" in Spanish Adolescents. *Apunts. Educació física i esports* 2010; 101: 19.
 21. CALE L. Self-report measures of children's physical activity: Recommendations for future development and a new alternative measure. *Health Educ J* 1994; 53: 439-453.
 22. SÁNCHEZ PT, HERNÁNDEZ BL. Validación de un cuestionario de actividad física habitual. *Apunts: Educación física y deportes* 1999; 58: 68-72.
 23. SALLIS JF, BUONO MJ, ROBY JJ, MICALE FG, NELSON JA. Seven-day recall and other physical activity self-reports in children and adolescents. *Med Sci Sports Exerc* 1993; 25: 99-108
 24. MARTÍNEZ-GÓMEZ D, ANDRES CALABRO M, WELK GJ, MARCOS A, VEIGA OL. Reliability and validity of a school recess physical activity recall in Spanish youth. *Pediatr Exerc Sci* 2010; 22: 218.
 25. MARTÍNEZ-GÓMEZ D, MARTÍNEZ-DE-HARO V, DEL-CAMPO J, ZAPATERA B, WELK GJ, VILLAGRA A et al. Validez de cuatro cuestionarios para valorar la actividad física en adolescentes españoles. *Gac Sanit* 2009; 23: 512-517.
 26. SERRA-MAJEM L. Crecimiento y desarrollo. Barcelona: Masson 2003; 113 p.
 27. MEREDITH MD, WELK GJ. Fitnessgram-Activitygram test administration manual. Champaign, IL: Human Kinetics 2007; 134 p.
 28. SALLIS JF, PATTERSON TL, BUONO MJ, NADER PR. Relation of cardiovascular fitness and physical activity to cardiovascular disease risk factors in children and adults. *Am J Epidemiol* 1988; 127: 933-941.
 29. SINGH AS, VIK FN, CHINAPAW MJ, UITDEWILLIGEN L, VERLOIGNE M, FERNÁNDEZ-ÁLVIRA JM, et al. Test-retest reliability and construct validity of the ENERGY-child questionnaire on energy balance-related behaviours and their potential determinants: the ENERGY-project. *Int J Behav Nutr Phys Act* 2011; 8: 136.
 30. VAN STRALEN MM, TE VELDE SJ, SINGH AS, DE BOURDEAUDHUIJ I, MARTENS MK, VAN DER SLUIS M, et al. European energy balance research to prevent excessive weight gain among youth (ENERGY) project: Design and methodology of the ENERGY cross-sectional survey. *BMC Public Health* 2011; 11: 1.
 31. BELTRAN EC, ADELL MN, HERETER JRV. Exercici i consum de tabac i alcohol dels escolars de Barcelona. *Gac Sanit* 1989; 3: 355-365.
 32. SUNNEGÅRDH J, BRATTEBY LE, SJÖLIN S. Physical Activity and Sports Involvement in 8 and 13Year Old Children in Sweden. *Acta Paediatr* 1985; 74: 904-912.
 33. HÄGSTROMER M, BERGMAN P, DE BOURDEAUDHUIJ I, ORTEGA FB, RUÍZ JR, MANIOS Y et al. Concurrent validity of a modified version of the International Physical Activity Questionnaire (IPAQ-A) in European adolescents: the HELENA study. *Int J Obes* 2008; 32 (Suppl 5): S42-48.
 34. CRAIG C, MARSHALL AL, SJOSTROM M, BAUMAN AE, BOOTH ML, AINSWORTH BE, et al., International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003; 35: 1381-1395.
 35. MARTÍNEZ-GÓMEZ D, MARTÍNEZ-DE-HARO V, POZO T, WELK GJ, VILLAGRA A, CALLE ME et al. Fiabilidad y validez del cuestionario de actividad física PAQ-A en adolescentes españoles. *Rev Esp Salud Publica* 2009; 83: 427-439.
 36. KOWALSKI CK, CROCKER PR, KOWALSKI NP. Convergent validity of the physical activity questionnaire for adolescents. *Pediatr Exerc Sci* 1997; 9: 342-352.

37. CANCELA JM, LAGO J, OUVIÑA L, AYÁN C. Validez del cuestionario de recuerdo de la actividad física realizada durante las 24h del día previo (PDPAR-24) en adolescentes españoles. *Nutr Hosp* 2015; 31: 1701-1707.
38. WESTON AT, PETOSA R, PATE RR. Validation of an instrument for measurement of physical activity in youth. *Med. Sci. Sports Exerc* 1997; 29: 138-143.
39. PROCHASKA JJ, SALLIS JF, LONG B. A physical activity screening measure for use with adolescents in primary care. *Arch Pediatr Adolesc Med* 2001; 155: 554-559.
40. GEISINGER KF. Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychol Assess* 1994; 6: 304.
41. NUNNALLY JC, BERNSTEIN LH. *Psychometric theory*, 3rd ed. New York: McGraw-Hill 1994; 752 p.
42. SHROUT PE, FLEISS JL. Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin* 1979; 86: 420.
43. DE VET HC, TERWEE CB, KNOL DL, BOUTER L. When to use agreement versus reliability measures. *J Clin Epidemiol* 2006; 59: 1033-1039.
44. KELLY P, FITZSIMONS C, BAKER G. Should we reframe how we think about physical activity and sedentary behaviour measurement? Validity and reliability reconsidered. *Int J Behav Nutr Phys Act* 2016; 13: 1.
45. POLS MA, PEETERS PH, KEMPER HC, GROBBEE D. Methodological aspects of physical activity assessment in epidemiological studies. *Eur J Epidemiol* 1998; 14: 63-70.
46. PATTERSON P. Reliability, validity, and methodological response to the assessment of physical activity via self-report. *Res Exerc Sport* 2000; 71 (2 Suppl): S15-20
47. RENNIE KL, WAREHAM NJ. The validation of physical activity instruments for measuring energy expenditure: problems and pitfalls. *Public Health Nutr* 1998; 1: 265-271.
48. FORSEN L, WAALER LOLAND N, VUILLEMIN A, CHINAPAW MJ, VAN POPEL MN, MOKKINK LB, et al., Self-administered physical activity questionnaires for elderly: a systematic review of measurement properties. *Sports Med* 2010; 40: 601-623.
49. VAN POPPEL MNN, CHINAPAW MJM, MOKKINK LB, VAN MECHELEN W. Physical activity questionnaires for adults: a systematic review of measurement properties. *Sports Med* 2010; 40: 565-600.
50. ALTMAN DG. *Practical statistics for medical research*. London: Chapman and Hall. 1991; 589 p.
51. GABRIEL KKP, MORROW JR, WOOLSEY ALT. Framework for physical activity as a complex and multidimensional behavior. *J Phys Act Health* 2010; (9 Suppl 1): S11-8.
52. BARANOWSKI T. Validity and reliability of self report measures of physical activity: an information-processing perspective. *Res Q Exerc Sport* 1988; 59: 314-327.
53. ROWE DA, MAHAR MT, RAEDEKE TD, LORE J. Measuring physical activity in children with pedometers: Reliability, reactivity, and replacement of missing data. *Pediatric Exercise Science* 2004; 16: 343-354.
54. ALTSCHULER A, PICCHI T, NELSON M, STERNFELD B. Physical activity questionnaire comprehension: lessons from cognitive interviews. *Med Sci Sports Exerc* 2009; 41: 336-343