

REVISIÓN SISTEMÁTICA

Recibido: 21 de abril de 2017
Aceptado: 3 de abril de 2018
Publicado: 14 de mayo de 2018

CUESTIONARIOS DE ESTIMACIÓN DE ACTIVIDAD FÍSICA: REVISIÓN SISTEMÁTICA Y ANÁLISIS DE SUS PROPIEDADES PSICOMÉTRICAS EN POBLACIÓN ESPAÑOLA MAYOR DE 60 AÑOS

Miguel A. Sánchez-Lastra (1), Iván Martínez-Lemos (2), José M. Cancela (1) y Carlos Ayán (2).

(1) Healthyfit Research Group. Departamento de Didácticas Especiales. Facultad de Ciencias de la Educación y el Deporte. Universidad de Vigo. Pontevedra. España.

(2) Departamento de Didácticas Especiales. Facultad de Ciencias de la Educación y el Deporte. Universidad de Vigo. Pontevedra. España.

Los autores declaran ausencia de conflictos de intereses.

RESUMEN

Fundamentos: Los cuestionarios de estimación de actividad física son herramientas ampliamente empleadas por los profesionales de la salud. Resulta esencial conocer la calidad y características de los mismos. Esta revisión sistemática tuvo como objetivo identificar y analizar las características psicométricas de los cuestionarios de realización de actividad física validados en población española mayor de 60 años, así como la calidad metodológica de los estudios de su validez y/o fiabilidad.

Métodos: Revisión sistemática de estudios donde se informe de la validez y/o fiabilidad de cuestionarios adaptados al castellano para estimar la actividad física realizada por población mayor de 60 años. Se estudiaron las características de la muestra, las propiedades psicométricas validez y fiabilidad test-retest, así como la calidad metodológica de los estudios, que se evaluó mediante la lista de comprobación *Qualitative Attributes and Measurement Properties of Physical Activity Questionnaires* (QAPAQ).

Resultados: Fueron encontrados y analizados un total de 8 cuestionarios, destacando en validez la *Versión Reducida en Español del cuestionario de actividad física en el tiempo libre de Minnesota* (VREM) (validez de constructo 95,5% con Kappa = 0,93) así como en fiabilidad (ICC=0,96), junto con el *Modified Baecke* (ICC = 0,92-0,96). Los criterios propuestos por el QAPAQ fueron superados en un 80% por 4 cuestionarios, siendo el *Yale Physical Activity Survey* (YPAS), el cuestionario que mayor puntuación obtuvo.

Conclusiones: Al objeto de estimar la cantidad de AF realizada por personas mayores de 60 años de nacionalidad española mediante cuestionarios, se aconseja el empleo, por este orden, de los instrumentos YPAS, VREM y *Modified Baecke*.

Palabras clave: Actividad Motora, Ejercicio físico, Terapia por Ejercicio, Encuestas y cuestionarios, Anciano, Validez de las pruebas, Reproducibilidad de resultados, Evaluación en Salud.

Correspondencia
Miguel A. Sánchez-Lastra
Facultad de Ciencias de la Educación y el Deporte
Universidad de Vigo
Campus A Xunqueira s/n
36005, Pontevedra
España
misanchez@uvigo.es

ABSTRACT

Physical activity questionnaires: a systematic review and analysis of their psychometric properties in Spanish population over 60 years old

Background: Questionnaires for estimating physical activity are measurement tools widely used by health professionals. Therefore, it is important to know their characteristics and quality. This systematic review aimed at identifying and analyzing the psychometric properties of the existing physical activity questionnaires validated in Spanish people older than 60 years of age, as well as to inform about the methodological quality of the studies performed concerning their validity and/or reliability.

Methods: A systematic review of the available scientific literature regarding studies focused on the validity and/or reliability of questionnaires used for estimating the prevalence of physical activity among Spanish populations over 60 years old. The Qualitative Attributes and Measurement Properties of Physical Activity Questionnaires checklist (QAPAQ), was used to evaluate the sample characteristics, the psychometric properties, the validity, the test-retest reliability and the methodological quality of each study identified.

Results: A total of 8 questionnaires were found and analyzed, highlighting in means of validity the short Spanish version (VREM) of the *Minnesota Leisure Time Physical Activity Questionnaire* (construct validity 95.5% with Kappa = 0.93) as well as in reliability (ICC=0.96), along with the *Modified Baecke* (ICC = 0.92-0.96). Four questionnaires met more than 80% of the criteria proposed by the QAPAQ. The YPAS was the questionnaire which reached the higher score.

Conclusions: When estimating the amount of physical activity performed by Spanish people older than 60 years of age by means of questionnaires, it is advisable to use (in this order) the YPAS, VREM and *Modified Baecke*.

Key words: Physical activity, Questionnaire, Energy expenditure, Elderly, Aged, Validity and reliability, Test-retest reliability, Geriatric Assessments.

Cita sugerida: Sánchez-Lastra MA, Martínez-Lemos I, Cancela JM, Ayán C. Cuestionarios de estimación de actividad física: revisión sistemática y análisis de sus propiedades psicométricas en población española mayor de 60 años. Rev Esp Salud Pública. 2018;92:14 de mayo e201805018.

INTRODUCCIÓN

Debido al proceso natural del envejecimiento, conforme avanza la edad, se produce un deterioro tanto a nivel de la función cognitiva^(1,2) como en la independencia funcional⁽³⁾. Esto último, relacionado en gran medida con el deterioro del equilibrio⁽⁴⁾, compromete el desempeño autónomo de las actividades de la vida diaria (AVDs), aumentando además el riesgo de caídas, las cuales producen fragilidad física y aumento de la dependencia, reduciendo la calidad de vida⁽⁵⁾. Este proceso suele desembocar en inactividad física, considerada como factor de riesgo para muchas enfermedades que comprometen la vida⁽⁶⁾.

Existen numerosas evidencias, recogidas en diferentes trabajos de revisión^(5,7), acerca de los beneficios que la práctica de ejercicio físico supone de cara al mantenimiento, mejora y ralentización del deterioro de no únicamente los parámetros de la independencia funcional, sino también a nivel de la función cognitiva. Para que la práctica de ejercicio físico conlleve realmente estos beneficios y no suponga riesgos adicionales, resulta necesario diseñar programas adecuados que estén correctamente contextualizados en las necesidades y características de este colectivo, asegurando una adecuada progresión que no solo minimice los riesgos de lesión, sino que favorezca la motivación y adherencia al programa, puntos esenciales para asegurar su efectividad⁽⁸⁾.

Para poder diseñar correctamente estos programas, es necesario el desarrollo de herramientas que evalúen el estado inicial de los individuos, siendo esencial que sean válidas y fiables para poder estudiar los efectos en la salud⁽⁹⁾. Los instrumentos más utilizados actualmente para la medición o estimación del nivel de condición física o actividad física (AF), se pueden clasificar en tres grupos principales: estimaciones mediante autoinformes o entrevistas (cuestionarios, escalas de puntuación, diarios), monitores de AF (acelerómetros, podómetros, monitores de frecuencia cardíaca) y métodos de medición directa del

gasto energético (agua doblemente marcada, calorimetría indirecta)⁽¹⁰⁾.

Los cuestionarios son generalmente más utilizados debido a su bajo coste y fácil administración, siendo además considerados particularmente útiles en grandes estudios epidemiológicos y longitudinales⁽¹⁰⁾.

Dado el uso generalizado que se hace de estas herramientas, se han publicado varios estudios de revisión acerca de sus propiedades psicométricas. Si bien informan principalmente sobre el grado de validez y fiabilidad de los cuestionarios, mayoritariamente son estudios centrados en la versión original de los mismos, profusamente diseñados para su aplicación en población de habla inglesa^(11,12). A este respecto, se debe destacar que para que los cuestionarios que han sido diseñados para un país y comunidad particular puedan ser utilizados en una población diana de distinta nacionalidad, deben ser traducidos mediante un procedimiento metodológico estandarizado denominado adaptación transcultural⁽¹³⁾. Este proceso toma especial relevancia debido a las peculiaridades culturales y medioambientales que definen a cada país⁽¹⁴⁾. Posteriormente, dicha versión debe ser validada en la población del país diana, demostrando unas adecuadas propiedades psicométricas⁽¹⁵⁾.

Hasta la fecha, únicamente existe un estudio de revisión, desarrollado por Guirao-Gorris *et al.*⁽¹⁶⁾, donde se informa detalladamente sobre las propiedades psicométricas de los cuestionarios, tanto de sus versiones originales como, en caso de su existencia, de sus traducciones y adaptaciones en poblaciones de diferentes países, incluyendo España. Esta revisión, por otro lado, no aporta información sobre la calidad metodológica del estudio de validación y adaptación para población española de los cuestionarios, al igual que no aporta información sobre la muestra en la que se estudian las propiedades psicométricas de dichos estudios. Sumado a ello, en este trabajo de revisión tampoco se recoge información acerca de los efectos suelo-techo de dichos estudios. Cabe destacar, de igual manera, que

esta revisión finalizó su búsqueda en 2007. Finalmente, a todo esto se debe añadir el hecho de que en las investigaciones en las que se revisan las características de instrumentos diseñados para monitorizar aspectos vinculados a la salud, es necesario aportar información sobre la calidad metodológica de las mismas, procedimiento que se realiza en base a criterios de consenso previamente establecidos a modo de lista de comprobación, como por ejemplo el COSMIN (Consensus-based standards for the Selection of health Measurement Instruments) “checklist”⁽¹⁷⁾. En el caso de los estudios de revisión centrados en las propiedades psicométricas que presentan los cuestionarios de valoración de la actividad física, se suele emplear una lista de comprobación específica creada a posteriori del COSMIN y conocida como QAPAQ (Qualitative Attributes and Measurement Properties of Physical Activity Questionnaires)⁽¹⁸⁾.

Debido a los puntos expuestos anteriormente, resulta necesaria una actualización de la temática, con un análisis en profundidad de los estudios de adaptación en población española, usando para ello una lista de comprobación específica.

Como consecuencia, se presenta el siguiente trabajo de revisión, que tuvo como objetivo identificar y analizar las características psicométricas de los cuestionarios de realización de AF validados en población española mayor de 60 años, así como la calidad metodológica de los estudios de acerca de su validez y/o fiabilidad.

MATERIAL Y MÉTODOS

Se realizó una revisión sistemática de estudios sobre cuestionarios de valoración de la AF en población española de adultos mayores de 60 años, publicados en inglés o español y sin acotación temporal. Se incluyeron los artículos que cumplían los siguientes criterios: a) estudio publicado en inglés o castellano; b) muestra formada específicamente por personas mayores de 60 años⁽¹⁹⁾ de nacionalidad española; c) utilizar una pregunta, serie de

preguntas o cuestionario validado; d) informar de la validez y/o fiabilidad de la pregunta, serie de pregunta o cuestionario empleado; e) en caso de utilizar cuestionarios que evalúen varios constructos además de la AF, informar sobre la validez y/o fiabilidad de las preguntas relacionadas específicamente con la misma.

Se consultaron las bases de datos electrónicas PubMed, Scopus, Sport Discus, Dialnet y The Cochrane Database of Systematic Reviews.

La estrategia de búsqueda se diseñó, en un primer momento, para encontrar el mayor número de estudios que tenían como objetivo comprobar la validez y/o fiabilidad de cuestionarios diseñados para valorar la realización de AF en población adulta mayor a nivel internacional. Posteriormente, se acotaron los resultados a los cuestionarios evaluados en dicho sector de la población, pero únicamente en España. Se realizó una búsqueda en las bases de datos anteriormente mencionadas con la combinación de palabras clave “Physical Activity”, “Questionnaire”, “Energy Expenditure” y “Older adults”, “Elderly” o “Older People”, seleccionando el último término en función del que obtuviese un mayor número de resultados en cada una de ellas. Posteriormente, se volvió a realizar una nueva búsqueda en las mismas bases de datos y añadiendo la palabra clave “Spain” a las anteriormente mencionadas, para revisar la existencia de nuevos estudios en España desde la primera prospección. La última búsqueda se realizó en diciembre de 2016. En la **tabla 1** se muestra una lista de todas las variables para las que se buscaron datos, anteriormente mencionadas, así como de sus definiciones.

Selección de estudios. Dos revisores examinaron independientemente la información incluida en el título y/o resumen de los resultados tras la primera búsqueda en las bases de datos (MS y CA). En base a las búsquedas, se realizó una primera selección donde se incluyeron todas las investigaciones que tuvieron como objetivo principal analizar la validez y/o fiabilidad de una pregunta, serie de preguntas

Tabla 1
Lista de variables para las que se buscaron datos y sus definiciones

Variable	Definición
Physical Activity	Estimación de la actividad física, la cual incluye al ejercicio físico, ampliando de esta forma los posibles resultados
Questionnaire	Herramienta empleada con el fin de que estimase la actividad física, o bien varios constructos incluyendo a ésta
Energy Expenditure	Estimación del gasto energético mediante la herramienta mencionada
Older Adults	Colectivo de edad superior a los 60 años
Elderly	Empleado en algunos casos de forma similar, para referirse a la población mayor de 60 o 65 años
Older people	Otro término empleado indiferentemente con los dos anteriores, para referirse al mismo colectivo de edad
Spain	Muestra de estudio española

o cuestionario cuya respuesta ofreciese información sobre el patrón de AF realizado por población española de edad superior a 60 años, identificándose la/s pregunta/s o el cuestionario empleado. No fueron incluidas las investigaciones que utilizarasen cuestionarios que evaluaran varios constructos además de la AF a no ser que incluyesen una pregunta o serie de preguntas sobre cuya validez/fiabilidad aportasen información por separado. Se localizó el texto completo de aquellos estudios que no presentasen en su título o resumen esta información. El grado de acuerdo inter-jueces se evaluó mediante el cálculo del índice Kappa de Cohen. Para los casos en que hubo falta de acuerdo, los revisores consultaron a un tercer autor (JC) y se llegó a un consenso. Posteriormente se llevó a cabo un segundo proceso de selección con el fin de localizar y revisar los estudios que, pese a juzgar por su título o resumen no tenían el objetivo anteriormente mencionado, se intuía que podrían aportar datos de relevancia al respecto. Finalmente, se revisaron los textos completos resultantes del segundo proceso de selección, para identificar el modo en el que se valoró el patrón de AF, excluyendo los artículos que no utilizasen una pregunta, serie de preguntas o cuestionario validado con ese objetivo, y que no informasen de la validez y/o fiabilidad de dichos ítems o herramientas. Todas las referencias bibliográficas de los

estudios incluidos fueron también revisadas para identificar potenciales investigaciones adecuadas para la inclusión.

Extracción de datos. Este paso lo llevaron a cabo los dos revisores anteriormente mencionados, los cuales recogieron los datos en una plantilla de forma independiente. La información extraída de los estudios incluidos se realizó mediante la lectura de los mismos. Para cada pregunta, serie de preguntas o cuestionario se recogieron 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 el nivel de AF para adultos mayores de 60 años⁽²⁰⁾. Para cada estudio se obtuvieron los siguientes datos: características de la muestra, variables evaluadas, grado de fiabilidad y/o validez informado, así como del procedimiento empleado para su valoración incluido el intervalo de tiempo entre medidas y la descripción de los instrumentos empleados (acelerómetros, otros cuestionarios y/o pruebas físicas). El grado de acuerdo inter-jueces se evaluó mediante el cálculo del índice Kappa de Cohen. Para los casos en que hubo falta de acuerdo, los revisores consultaron a un tercer autor (JC) y se llegó a un consenso.

Evaluación del riesgo de sesgo. Para la evaluación del riesgo de sesgo, se utilizó la

versión modificada del checklist de Downs y Black⁽²¹⁾ llevada a cabo por Prince et al.⁽²²⁾. La escala original se desarrolló para ser empleada en estudios observacionales y consiste en 27 ítems con una puntuación máxima de 32 puntos. La versión modificada, se realizó con el objetivo de adaptar este instrumento con el fin de evaluar el riesgo de sesgo en una revisión sobre medidas directas y auto-informes para evaluar la actividad física en adultos. Debido a ello, se consideró óptima para su utilización en el presente estudio. La versión modificada utilizada, comprende los ítems 1-4, 6, 7, 9-13, 16-18 y 20 de la lista original, con una puntuación máxima de 15, indicando mayor calidad a mayor puntuación. Dos autores llevaron a cabo dicha evaluación independientemente (MS y CA). En caso de diferencias, se llegó a un acuerdo mediante la discusión de los mismos. Debido a la heterogeneidad de los estudios y al reducido número de los mismos, no se consideró posible un análisis del sesgo que pudiese afectar a la evidencia acumulativa.

Propiedades psicométricas. Se recogió la información aportada por cada uno de los estudios acerca de las características de la muestra (número, sexo y estado de salud diana). Se incluyó información acerca de la fiabilidad test-retest de los estudios que la hubiesen analizado, incluyendo los días entre las diferentes administraciones. Dada la ausencia de un patrón oro que permita estimar la validez de criterio de los cuestionarios diseñados para estimar la cantidad de AF realizada⁽²²⁾, en la presente revisión se identificó como validez de constructo todos aquellos diseños metodológicos que mediante el empleo de acelerómetros, podómetros, pruebas de condición física o parámetros antropométricos (IMC) fueron desarrollados en las investigaciones localizadas con el fin de determinar la validez de los cuestionarios objeto de estudio^(23,24,25).

Calidad metodológica. Se evaluó la calidad metodológica de los estudios incluidos mediante la lista de comprobación QAPAQ⁽¹⁸⁾. 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 los siguientes 10 criterios, que fueron calificados como ausentes o presentes en función de si se aportaba o no la información correspondiente: Descripción completa de la muestra, diseño del estudio, procedimiento de administración, error de medida, coeficientes de fiabilidad, validez aparente, validez de contenido, efectos suelo-techo, validez de constructo y sensibilidad. La puntuación máxima por estudio fue, por tanto, de 10.

RESULTADOS

La búsqueda inicial permitió identificar un total de 662 estudios y la segunda 5 artículos más recientes. De ellos, se seleccionaron 96 textos para su lectura completa, tras la cual se excluyeron 89 (figura 1). De esta forma, se identificaron 7 estudios que informaron sobre la validez/fiabilidad de 8 cuestionarios redactados en lengua castellana con el objetivo de cuantificar la AF realizada por adultos mayores españoles (tabla 2). Se alcanzó un aceptable grado de acuerdo entre revisores, tanto en el proceso de selección de estudios como en el de extracción de datos (Kappa= 0.61 (0.52-0.68) y Kappa= 0.72 (0.35-1.0)).

Siguiendo la estructura del QAPAQ, en la tabla 3 se muestra la descripción de los atributos cualitativos de los cuestionarios. Como se puede observar, algunos de ellos fueron originalmente diseñados en población adulta mayor y otros son producto de adaptaciones de cuestionarios creados originalmente para su administración en población adulta. Tres de los cuestionarios analizados pertenecen al primer grupo: Yale Physical Activity Survey (YPAS)^(26,27,28), Modified Baecke Physical Activity Questionnaire (Modified Baecke PAQ)^(29,30) y

Figura 1
Diagrama de flujo del procedimiento de selección de las investigaciones localizadas en el estudio

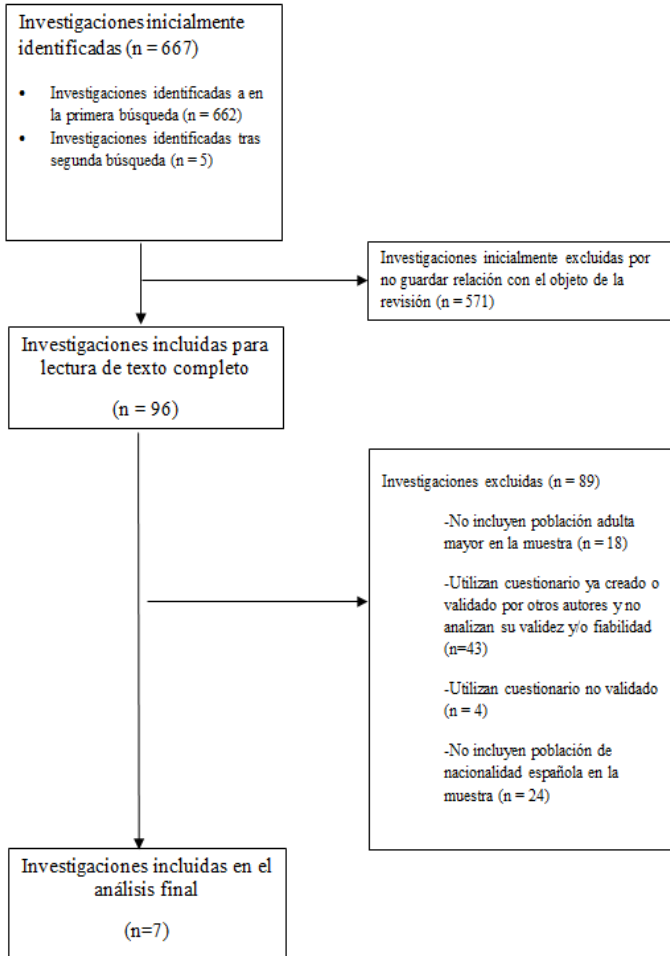


Tabla 2					
Lista de los cuestionarios en castellano para estimar la actividad física analizados					
Instrumento y referencia original	Estudios de validación existentes: Estado de salud diana	País original de administración^b	Lengua^e	Adaptación cultural^f	Cumplimiento de recomendaciones^g
YPAS ²⁸	Sin patologías específicas ²⁶ y Pacientes con EPOC ²⁷	EEUU	Inglés	Si	Si
IPAQ-SF ³³	Supervivientes de cáncer ³⁴	Varios ^c	Varios ^c	Si	Si
GPAQ ³⁵	Supervivientes de cáncer ³⁴	Varios ^d	Varios ^d	NI	Si
Modified Baecke PAQ ³⁰	Pacientes con EPOC ²⁹	Holanda	Inglés	Si	No
VREM ^{*34}	Sin patologías específicas ³⁵	EEUU	Inglés	No	Si
CHAMPS ³²	Sin patologías específicas ³¹	EEUU	Inglés	Si	Sí
LTPAI ³⁷	Sin patologías específicas ³⁸	Suecia	Inglés	Si	Si
PAWHI ³⁹	Sin patologías específicas ³⁸	Suecia	Sueco	Si	Si

*Validación de una adaptación propia del Minnesota Leisure Time Physical Activity Questionnaire; ^bPaís original de validación; ^cEl 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); ^dEl estudio de validez y fiabilidad se realizó en 9 países (Bangladesh, Brasil, China, Etiopía, India, Indonesia, Japón, Portugal y Sudáfrica); ^eLengua original de validación; ^fEn el procedimiento de traducción se realizó retro-traducción; ^gCapacidad para discriminar el grado de cumplimiento de las recomendaciones de consenso sobre el nivel de actividad física saludable para adultos mayores; NI: No hay información.

el Community Healthy Activities Model program for Seniors Questionnaire (CHAMPS)^(31,32). En el segundo grupo están encuadrados los 5 restantes: La versión corta del International Physical Activity Questionnaire (IPAQ)⁽³³⁾ denominada IPAQ - Short Form (IPAQ-SF)⁽³⁴⁾, el Global Physical Activity Questionnaire (GPAQ)^(34,35), una versión reducida del Minnesota Leisure Time Activity Questionnaire (MLTAQ)⁽³⁶⁾ denominada Versión Reducida en Español del cuestionario de actividad física en el tiempo libre de Minnesota (VREM)⁽³⁷⁾ y los cuestionarios adaptados Leisure Time Physical Activity Instrument (LTPAI)⁽³⁸⁾ y Physical Activity at Home and Work Instrument (PAWHI)⁽³⁸⁾, originariamente diseñados para población con fibromialgia⁽³⁹⁾. Ninguno de los cuestionarios fue creado originalmente para su administración en población adulta mayor en lengua castellana, por lo que todos fueron adaptados a partir de su versión original en diferentes lenguas. En referencia al proceso de adaptación del idioma, seis de ellos informaron de un método estándar de adaptación cultural⁽¹³⁾. En la misma tabla se puede apreciar que cinco de los instrumentos se administran mediante entrevista y tres de ellos son autoinformes.

En la **tabla 4** se resumen las principales características de las investigaciones que analizaron la validez y/o fiabilidad de los cuestionarios. En cinco estudios, la muestra estuvo conformada por población comunitaria^(26,31,34,37,38) y en dos por población clínica, concretamente pacientes con Enfermedad Pulmonar Obstructiva Crónica (EPOC)^(27,29). El único cuestionario validado tanto en población comunitaria como clínica fue el YPAS^(26,27). Se informó de la validez de constructo en todos los cuestionarios, utilizándose para ello la comparación con acelerómetros u otros cuestionarios validados en seis de los estudios localizados, cinco en población comunitaria (citas) y uno en clínica⁽²⁷⁾. A este respecto, destacó el VREM (validez de constructo 95,5% con Kappa = 0,93), que fue comparado con otro cuestionario validado, así como los ítems relacionados con una AF vigorosa en los

cuestionarios IPAQ-SF y GPAQ, con un coeficiente de correlación lineal de Pearson (r) de 0,73 y 0,98 respectivamente, comparados con un acelerómetro. En cuatro investigaciones, la validez fue analizada mediante la comparación con pruebas físicas y/o parámetros de la composición corporal. Este procedimiento se empleó tanto en población comunitaria^(31,38) como clínica^(27,29). Los resultados indicaron que los cuestionarios Modified Baecke PAQ ($r = 0,54$)⁽²⁹⁾ y el YPAS⁽²⁷⁾ (coeficiente de correlación de Spearman (ρ) de 0,40) mostraron los valores más altos, comparados con el test de la marcha de seis minutos (6MWT), seguidos por el CHAMPS en los ítems gasto energético total por semana ($r=0,498$) y minutos totales por semana ($r = 0,519$), comparado con el Índice de Barthel (IB). Se informó de la fiabilidad test-retest de cinco cuestionarios, tres con muestra comunitaria y dos en clínica, de los cuales el Modified Baecke PAQ (ICC = 0,92-0,96) el VREM (ICC = 0,96) y el YPAS⁽²⁶⁾ (ICC = 0,65 para el ítem tiempo total e ICC = 0,66 en el gasto energético diario) alcanzaron los valores más altos.

Se pueden observar en la **tabla 5**, finalmente, los resultados de la aplicación del instrumento QAPAQ y sus criterios de análisis para evaluar las propiedades de estimación de los instrumentos analizados. En referencia a la cuantificación del error de medida, únicamente cuatro de los instrumentos informaron del mismo. Cabe destacar que, como se muestra en dicha tabla, si bien todos superaron un grado de cumplimiento $\geq 50\%$ de los criterios de calidad establecidos, ninguno de los estudios de los cuestionarios analizados cumple el 100% de los mismos.

En el presente trabajo de revisión, se evaluó el riesgo de sesgo en todos los estudios ($n=7$). El rango de los resultados en los ítems de la herramienta modificada de Downs y Black⁽²¹⁾ fue de 12 a 14 (el máximo era 15), con una media de $13 \pm 0,58$. Los resultados de la evaluación del riesgo de sesgo indican que todos los estudios tuvieron una alta calidad, basada en un punto de corte de superior a 12 sobre los

Tabla 3
Descripción de los atributos cualitativos de los cuestionarios validados al castellano para medir actividad física en adultos mayores (n=8)

Cuestionario	Constructo	Población diana: rango de edad (años)	Escenario de práctica	Administración	Período de recuerdo	Formato			Interpretabilidad: Unidad de medida
						Ítems	Lista de actividades	Dimensiones	
YPAS	Gasto energético	>60	Doméstico, jardín, cuidado de personas, ejercicio y ocio	Entrevista	Típica semana del último mes	40	Sí	T/F/D/I/Cs	Horas · sem ⁻¹ ; kJ · día ⁻¹
IPAQ-SF	Gasto energético	15-69	Doméstico, trabajo, transportes y ocio	Autoinforme	Últimos 7 días	9	Sí	T/F/D/I/Cs	METs · min ⁻¹ · día ⁻¹ ; Min · día ⁻¹
GPAQ	Gasto energético	>18	Trabajo, desplazamiento y ocio	Entrevista	Semana típica y día típico	16	Sí	T/F/D/I/Cs	Min · sem ⁻¹ ; METs – min/sem
Modified Baecke PAQ	AF habitual	> 60	Doméstico, deporte y ocio	Entrevista	Habitual	14	Sí	T/F/D/I/Cs	Escala de puntuación: <9 = sedentario, 9 a 16 = sedentarios moderados, <16 = activos.
VREM	Gasto energético	18-60	Caminar, trabajar en el huerto, hacer deporte o bailar, subir escaleras, ir a comprar a pie, limpiar la casa	Entrevista	Última semana o semana habitual, último mes o mes habitual y meses por año	6	Sí	T/F/D/I	METs - min/14 días
CHAMPS	Gasto energético	>60	Doméstico, ejercicio y ocio	Entrevista	Semana típica del último mes	31	Sí	T/F/D	Min · sem ⁻¹ ; Kcal · sem ⁻¹
LTPAI	AF habitual	30-65	Ejercicio y ocio	Autoinforme	Semana típica del último mes	4	No	T/D/I/Cs	Horas · sem ⁻¹
PAWHI	AF habitual	30-65	Doméstico y trabajo	Autoinforme	Semana típica del último mes	7	No	T/D/I/Cs	Horas · sem ⁻¹

Cs: Conducta sedentaria; D: Duración; F: Frecuencia; I: Intensidad; MET: Equivalente metabólico, unidad empleada para calcular el gasto de energía reflejado en el consumo de oxígeno de cualquier actividad física; Min: Minutos; NI: no informado; Sem: Semana; T: Tipo.

Tabla 4
Análisis de validez y fiabilidad de los cuestionarios incluidos en la selección final (n=9)

Cuestionario y estudio	Muestra de Estudio Tamaño (edad media \pm DS); % género femenino; estado de salud diana	Ítems evaluados	Resultados			
			Validez	Intervalo (días)	Fiabilidad (test-retest)	
YPAS ²⁶	n = 108 (69,8 \pm 7,1 años hombres / 67,7 \pm 6,4 años mujeres); 64,81% Mujeres; Sin patologías específicas	Tiempo total (hr/sem)	Ac*: r = 0,20*; IMC: r = -0,02; Peso (kg): r = -0,26*; GC: r = 0,13	7	ICC = 0,66 ^c	
		Gasto energético (kJ/día)	Ac*: r = 0,23*; IMC: -0,12; Peso (kg): r = -0,25*; GC: r = -0,11		ICC = 0,65 ^c	
		Índices de actividades	Sumatoria (unidades totales)		Ac*: r = 0,24*; IMC: r = -0,22*; Peso (kg): r = -0,12; GC: r = -0,27 ^b	ICC = 0,31 ^b
			Vigoroso (unidades/mes)		Ac*: r = -0,01; IMC: r = -0,29; Peso (kg): r = -0,12; GC: r = -0,08	ICC = 0,20 ^a
			Paseo de ocio (unidades/mes)		Ac*: r = 0,24*; IMC: r = -0,05; Peso (kg): r = -0,03; GC: r = 0,3 ^a	ICC = 0,33 ^c
			Moviéndose (unidades/mes)		Ac*: r = 0,31*; IMC: r = -0,04; Peso (kg): r = -0,14; GC: r = -0,10	ICC = 0,19 ^a
			Estar de pie (unidades/mes)		Ac*: r = 0,19; IMC: r = 0,04; Peso (kg): r = -0,11; GC: r = -0,07	ICC = 0,12
Sentado (unidades/mes)	Ac*: r = -0,06; IMC: r = 0,09; Peso (kg): r = 0,20*; GC: r = -0,05	ICC = 0,29 ^b				
IPAQ-SF ³⁴	n = 204 (54 \pm 11 años); 64% Mujeres; Supervivientes de Cáncer	AF Vigorosa (Min \cdot sem ⁻¹)	Ac***: r = 0,98c-; CI (0,97-0,99)	NI	NI	
		AF Moderada (Min \cdot sem ⁻¹)	Ac***: r = -0,04; CI (-0,19-0,11)			
		Inactividad física (Min \cdot sem ⁻¹)	Ac***: r = 0,07; CI (-0,08-0,22)			
GPAQ ³⁴	n = 204 (54 \pm 11 años); 64% Mujeres; Supervivientes de Cáncer	AF Vigorosa (Min \cdot sem ⁻¹)	Ac***: r = 0,73c-; CI (0,65-0,79)	NI	NI	
		AF Moderada (Min \cdot sem ⁻¹)	Ac***: r = -0,03; CI (-0,18-0,12)			
		Inactividad física (Min \cdot sem ⁻¹)	Ac***: r = 0,07*; CI (0,02-0,31)			
Modified Baecke PAQ ²⁹	n = 55 (66 \pm 8 años); 5% mujeres; Pacientes con EPOC	Puntuación total	FEV1; rho = 0,31*; CdV: rho = -0,45*; 6MWT: rho = 0,54 ^a	>7; <31	ICC = 0,96	
		Actividades domésticas	-		ICC = 0,93	
		Ocio	-		ICC = 0,92	
		Deporte	-		ICC = 0,95	

Ac*: Acelerómetro Caltrac; Ac***: Acelerómetro ActiGraph GT3X; Ac***: Acelerómetro SenseWear Pro2 Armband; CI: Intervalo de confianza 95%; EPOC: Enfermedad pulmonar obstructiva crónica; 5CSTS:: Prueba 5-Chair Sit-to-Stand Test; FEV1: Volumen espiratorio forzado en el primer segundo; GC: Porcentaje de grasa corporal; ICC: Coeficiente de correlación intraclass; IMC: Índice de masa corporal (kg/m2); Kappa: Coeficiente kappa de Cohen; MET: Equivalente metabólico, unidad empleada para calcular el gasto de energía reflejado en el consumo de oxígeno de cualquier actividad física; NI (no informada); r: Coeficiente de correlación de Pearson; rho: Coeficiente de Correlación de Spearman; Sem: semana; 6MWT: Test de la marcha de 6 minutos; *p<0,05 ; ^bp<0,01; ^cp<0,001; + Cuestionario YPAS; ++ Cuestionario de Actividad Física en el tiempo libre de Minnesota.

Tabla 4. Cont.
Análisis de validez y fiabilidad de los cuestionarios incluidos en la selección final (n=9)

Cuestionario y estudio	Muestra de Estudio Tamaño (edad media ± DS); % género femenino; estado de salud diana	Ítems evaluados	Resultados		
			Validez	Intervalo (días)	Fiabilidad (test-retest)
VREM ³⁷	n = 200 (65,88 años); 49% Mujeres; Sin patologías específicas	METS-min/14 días	Cuestionario++; Validez de criterio = 95,5%; Kappa = 0,93; CI (0,84-1,00).	15	ICC = 0,96 (0,95-0,98)
YPAS ²⁷	N = 161 (70 ± 8 años); 41,61% mujeres; Pacientes con EPOC	Tiempo en actividad ≥ 1,4 MET (h/día)	Ac*** (Número de pasos); rho = 0,34; Ac*** (Tiempo en actividad ≥ 1,4 MET (h/día): rho = 0,38; Ac*** (MET en actividad ≥ 1,4 (MET-h/día): rho = 0,35; Ac*** (GE en actividad ≥ 1,4 METs): rho = 0,36; 6MWT: rho = 0,37 ^c	NI	NI
		MET en actividad ≥ 1,4 (MET-h/día)	Ac*** (Número de pasos); rho = 0,38; Ac*** (Tiempo en actividad ≥ 1,4 MET (h/día): rho = 0,41; Ac*** (MET en actividad ≥ 1,4 (MET-h/día): rho = 0,38; Ac*** (GE en actividad ≥ 1,4 METs): rho = 0,40; 6MWT: rho = 0,37 ^c		
		GE en actividad ≥ 1,4 METs (Kcal/día)	Ac*** (Número de pasos); rho = 0,32c; Ac*** (Tiempo en actividad ≥ 1,4 MET (h/día): rho = 0,32; Ac*** (MET en actividad ≥ 1,4 (MET-h/día): rho = 0,29; Ac*** (GE en actividad ≥ 1,4 METs): rho = 0,37; 6MWT: rho = 0,33 ^c		
		Puntuación total	Ac*** (Número de pasos); rho = 0,52; Ac*** (Tiempo en actividad ≥ 1,4 MET (h/día): rho = 0,38; Ac*** (MET en actividad ≥ 1,4 (MET-h/día): rho = 0,42; Ac*** (GE en actividad ≥ 1,4 METs): rho = 0,43; 6MWT: rho = 0,40 ^c		
CHAMPS ³¹	n = 52 (82,40 ± 8,26 años); 72,2% Mujeres; Sin patologías específicas	GE/sem en AF de intensidad moderada o mayor	Ac*: r = 0, 253 ^a ; IB: r = 0,383 ^a ; 5CSTS: r = -0,274	NI	NI
		Min/sem en AF de intensidad moderada o mayor	Ac*: r = 0,528 ^a ; IB: r = 0,313 ^a ; 5CSTS: r = -0,236a		
		GE/sem total	Ac*: r = 0,287 ^a ; IB: r = 0,498 ^b ; 5CSTS: r = -0,170		
		Min/sem total	Ac*: r = 0,385 ^a ; IB: r = 0,519 ^b ; 5CSTS: r = -0,132		
LTPAI ³⁸	n = 44 (68,11 ± 5,38 años); 100% mujeres; Sin patologías específicas	Puntuación total	Cuestionario++; rho = 0,28 ^a ; 6MWT: rho = 0,002	15	ICC = 0,38 (0,14-0,66)
		Ligero	-		ICC = 0,41 (0,008-0,67)
		Moderado	-		ICC = -0,23 (-1,25-0,33)
		Vigoroso	-		ICC = 0,06 (-0,72-0,48)
PAHWI ³⁸	n = 44 (68,11 ± 5,38 años); 100% mujeres; Sin patologías específicas	Puntuación total	NI	15	ICC = 0,55 (0,18-0,75)
		Ligero			ICC = 0,40 (-0,10-0,67)
		Moderado			ICC = 0,17 (-0,77-0,61)
		Vigoroso			NI

Ac*: Acelerómetro Caltrac; Ac**: Acelerómetro ActiGraph GT3X; Ac***: Acelerómetro SenseWear Pro2 Armband; CI: Intervalo de confianza 95%; EPOC: Enfermedad pulmonar obstructiva crónica; 5CSTS:: Prueba 5-Chair Sit-to-Stand Test; FEV1: Volumen espiratorio forzado en el primer segundo; GC: Porcentaje de grasa corporal; ICC: Coeficiente de correlación intraclass; IMC: Índice de masa corporal (kg/m²); Kappa: Coeficiente kappa de Cohen; MET: Equivalente metabólico, unidad empleada para calcular el gasto de energía reflejado en el consumo de oxígeno de cualquier actividad física; NI (no informada); r: Coeficiente de correlación de Pearson; rho: Coeficiente de Correlación de Spearman; Sem: semana; 6MWT: Test de la marcha de 6 minutos; *p<0,05 ; **p<0,01; ***p<0,001; + Cuestionario YPAS; ++ Cuestionario de Actividad Física en el tiempo libre de Minnesota.

Tabla 5
Evaluación de las propiedades de medida de los cuestionarios de valoración de AF en población española de adultos mayores (n=8)

Cuestionario y estudio	General			Fiabilidad			Validez			Sensibilidad ^j	Puntos sobre el total
	Población de estudio ^a	Diseño del estudio ^b	Adm	Error de medida ^d	Test-Retest ^e	Aparente ^f	Contenido ^g	Efecto suelo-techo ^h	Constructo ⁱ		
YPAS ²⁶	+	+	+	+	+	+	+	-	+	+	9/10
IPAQ-SF ³⁴	+	+	+	+	-	+	+	-	+	+	8/10
GPAQ3 ⁴	+	+	+	+	-	+	+	-	+	+	8/10
Mod Baecke PAQ ²⁹	+	+	+	-	+	+	+	+	+	-	8/10
VREM ³⁵	+	+	+	+	+	+	-	-	+	-	7/10
YPAS ²⁷	+	+	+	-	-	+	+	-	+	-	6/10
CHAMPS ³¹	+	+	+	-	-	+	+	-	+	-	6/10
LTPAI ³⁸	+	+	+	-	+	-	-	-	+	-	5/10
PAHWI ³⁸	+	+	+	-	+	-	-	-	+	-	5/10

^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 Constructo (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).

15 puntos⁽²²⁾. Como se ha indicado anteriormente, no se realizó un análisis del riesgo de sesgo entre los estudios debido a su reducido número y heterogeneidad.

DISCUSIÓN

El análisis de las investigaciones seleccionadas, demuestra que no todos los cuestionarios de recuerdo de realización de AF aplicables en población adulta mayor de 60 años española llevan a cabo una evaluación conveniente de sus propiedades psicométricas. De este modo, del total de los 7 estudios se informó en 7 de ellos, en el caso del YPAS dos veces^(26,27), sobre su validez y en 5 cuestionarios sobre su fiabilidad test-retest.

En referencia al análisis de la fiabilidad, únicamente dos cuestionarios que podrían calificarse con una alta fiabilidad⁽⁴⁰⁾, el Modified Baecke PAQ y el VREM, mostraron valores superiores al estándar mínimo⁽⁴¹⁾.

Además de esta limitación, únicamente 4 estudios incluyeron información acerca de la cuantificación del error de medida. Esta propiedad permite informar del error aleatorio en la puntuación del cuestionario por medio de diferentes parámetros como el límite del acuerdo, el mínimo cambio detectable y/o el tamaño medio del error⁽⁴²⁾. Dicho error de medida toma especial relevancia en estudios que pretenden medir el efecto de una intervención⁽⁴³⁾.

Con respecto al análisis de validez, el mayor nivel de evidencia, denominada validez de criterio, se obtendría comparando el cuestionario con un patrón oro, un instrumento con una perfecta validez y fiabilidad, que en el caso de la medición de la AF no existe^(22,44,45). Si bien se consideraba el agua doblemente marcada como el patrón oro para estimar el gasto energético total diario⁽⁴⁶⁾, este método no cumple al completo los requisitos para tal denominación, pues dicho gasto no es causado únicamente por la AF, sino también por el índice metabólico basal y el efecto térmico de los alimentos⁽²²⁾. De igual manera,

tampoco permite diferenciar entre el tipo, frecuencia y duración de las actividades⁽⁴³⁾. Debido a la ausencia, por tanto, del denominado patrón oro, podríamos informar acerca de la validez de constructo, mediante la comparación del cuestionario con otro instrumento validado que mida constructos fuertemente relacionados, como los acelerómetros u otro cuestionario previamente validado⁽²²⁾. De igual manera, diversos autores utilizan las pruebas de condición física y las mediciones antropométricas y de la composición corporal para el estudio de la validez de constructo⁽²³⁾, aunque de forma más indirecta y por ello, normalmente, complementaria con el uso de acelerómetros^(25,47).

Si bien no hay un firme consenso sobre los valores de correlación que indicarían si la validez es o no adecuada, diferentes autores han considerado como valor mínimo estándar un coeficiente de correlación de Pearson, $r \geq 0,50$ ^(23,48). De esta manera, únicamente el VREM habría obtenido unos resultados por encima de dicho valor en referencia a las puntuaciones totales. En el caso del GPAQ y el IPAQ-SF, sólo los ítems relacionados con una AF de intensidad vigorosa lo superan.

Con respecto al tamaño de la muestra para evaluar las propiedades de medida, en general fue superior al tamaño muestral mínimo aconsejable ($n \geq 50$)⁽⁴⁹⁾, excepto en el estudio de los cuestionarios LTPAI y PAWHI, donde no se cumplió dicha recomendación.

La comparación de los valores de las propiedades psicométricas de los cuestionarios entre su versión original y la adaptada en castellano ofrece resultados dispares. En relación al análisis de la validez, el valor obtenido con la versión española del YPAS en 108 sujetos sin patologías específicas⁽²⁶⁾ es superior al obtenido en su versión original⁽²⁸⁾ en una muestra de 25 adultos mayores de 60 años. En el caso del CHAMPS, los resultados son ligeramente inferiores en la versión en castellano⁽³¹⁾ frente a su versión original⁽⁵⁰⁾, en muestras de $n = 52$ y $n = 87$, respectivamente. En esta línea, la versión en castellano⁽³⁸⁾ del LTPAI, en

un $n = 44$, obtuvo unos valores ampliamente inferiores a los de su versión original⁽³⁹⁾ con $n = 37$. No se pudo establecer comparaciones en los otros casos. Por otro lado, en relación a la comparación de la fiabilidad, el valor obtenido con la versión española del YPAS⁽²⁸⁾ en los ítems tiempo total y gasto energético total diario, fueron mayores que en su versión original⁽²⁸⁾, no así en el caso de la sumatoria de los índices de actividades, donde fue inferior. En el caso del LTPAI, el valor obtenido con la versión en español⁽³⁸⁾ es muy inferior al obtenido en su versión original⁽³⁹⁾, guardando similitud con la versión en español del PAWHI⁽³⁸⁾, que también consiguió unos resultados inferiores al original⁽³⁹⁾.

En lo referente a los atributos cualitativos de los cuestionarios, se constató una gran disparidad. Así, de los cuatro escenarios principales de práctica de AF (ocio, trabajo, doméstico y deporte), la mayoría de instrumentos tienen en cuenta hasta 3 de ellos, encontrando que únicamente en el LTPAI y el PAWHI se evalúan solamente dos escenarios; ejercicio y ocio en el primero, doméstico y trabajo en el segundo. La ausencia de escenarios recogidos en el cuestionario reduce la precisión con la que los mismos pueden estimar la AF realizada⁽⁵¹⁾.

Al margen de los atributos cualitativos, se debe tener en cuenta que, si bien aparece reflejado en diferentes estudios que los adultos mayores son la población con una edad superior o igual a 65 años, en esta revisión sistemática se optó por utilizar como punto de corte de edad para considerar la muestra como adultos mayores, los 60 años, siguiendo la utilización de dicho término por la Organización Mundial de la Salud⁽⁵²⁾ y permitiendo así una recuperación y análisis de un mayor número de estudios.

En relación a este punto, la edad, una importante debilidad observada en estos instrumentos es que tienen algunas limitaciones asociadas con el sesgo de recuerdo, donde suelen sobreestimar el tiempo en actividades de intensidad vigorosa y subestimar el tiempo

de actividades que duran menos de 10 minutos o con un nivel de esfuerzo inferior a la marcha enérgica⁽⁵²⁾. El hecho de que varios de los instrumentos aquí incluidos tengan un tiempo de recuerdo superior a 7 días a excepción del YPAS⁽²⁶⁾, acentúa esta problemática, principalmente debido a la posible presencia de deterioro en la memoria a corto plazo, acrecentado conforme la edad aumenta⁽⁷⁾. Sumada a la anterior, otra debilidad está relacionada con la ausencia de una definición concreta que permita diferenciar los diferentes tipos de intensidad de la AF realizada, pues el factor subjetivo de cada respuesta podría alterar también en cierta medida los resultados.

En contraposición, una característica positiva observada fue la posibilidad de que la información recogida por el cuestionario en forma de puntuación pueda ser expresada en unidades estándar de gasto energético, Estimación de la Tasa Metabólica por minuto al día (METs·min⁻¹·día⁻¹); Kilocalorías por semana (Kcal·sem⁻¹), o que permitan su conversión a las mismas desde unidades de tiempo como la duración o la frecuencia. Esto resulta de gran interés pues permite identificar de forma rápida y directa, a partir del resultado obtenido en el cuestionario por parte del encuestado, si éste cumple las recomendaciones de AF publicadas para este grupo poblacional, los adultos mayores⁽²⁰⁾.

Los resultados obtenidos en esta revisión indican por un lado que, en relación a las propiedades psicométricas estudiadas, los cuestionarios con un mejor comportamiento de la validez han sido el VREM y el Modified Baecke PAQ, al igual que en relación a la fiabilidad, superando en ambos casos los estándares mínimos establecidos para dichas propiedades.

Un aspecto a destacar es el hecho de que algunos de los cuestionarios aquí presentados y analizados (por ej., el Modified Baecke PAQ y el YPAS), fueron originariamente creados y validados en población comunitaria, si bien los estudios de adaptación cultural incluidos en esta revisión fueron realizados en pobla-

ción clínica. Por otro lado, también se observó la situación inversa, puesto que los cuestionarios PAHWI y el LTPAI fueron originalmente creados para estimar la cantidad de AF realizada por personas con fibromialgia, mientras que su adaptación al castellano se realizó con población sana. De todos modos, gran parte de los cuestionarios localizados en la presente revisión fueron diseñados para ser aplicados en población comunitaria, y las muestras incluidas en los estudios de validación al castellano presentaron esta característica. En relación al único cuestionario validado tanto en población comunitaria como clínica, el YPAS, se debe mencionar que mostró un mayor grado de validez cuando se administró en este último colectivo.

En relación a los criterios de calidad metodológica de acuerdo con el instrumento QAPAQ, se debe remarcar que el YPAS(26) cumple el 90% de dichos criterios, mientras que el GPAQ⁽³⁴⁾, el IPAQ-SF⁽³⁴⁾ y el Modified Baecke PAQ⁽²⁹⁾ cumplen con el 80% de los mismos, si bien este último cuestionario sólo parece de utilidad en estudios transversales. El resto de instrumentos analizados se situaron por debajo de este porcentaje. En el caso del VREM, sólo se cumplen el 60% de los mismos.

No se ha encontrado, por tanto, ningún instrumento que cumpla los requisitos metodológicos del QAPAQ de forma completa o superando el 70% de los mismos, que alcance a su vez unos valores de validez y fiabilidad adecuados.

A la vista de los resultados aquí presentados, al objeto de estimar la cantidad de AF realizada por personas mayores de 60 años de nacionalidad española mediante cuestionarios, se aconseja el empleo, por este orden, de los instrumentos YPAS, VREM y Modified Baecke PAQ.

BIBLIOGRAFÍA

1. Sofi F, Valecchi D, Bacci D, Abbate R, Gensini GF, Casini A, et al. Physical activity and risk of cognitive decline: a meta-analysis of prospective studies. *J Intern Med*. 2011;269(1):107-117.

2. Aichberger MC, Busch MA, Reischies FM, Ströhle A, Heinz A, Rapp MA. Effect of physical inactivity on cognitive performance after 2.5 years of follow-up: Longitudinal results from the Survey of Health, Ageing, and Retirement (SHARE). *GeroPsych (Bern)*. 2010;23(1):7-15.

3. Martínez-Gomez D, Guallar-Castillon P, Rodríguez-Artalejo F. Sitting Time and Mortality in Older Adults With Disability: A National Cohort Study. *J Am Med Dir Assoc*. 2016;17(10):960.e15-20.

4. Hausdorff JM, Rios DA, Edelberg HK. Gait variability and fall risk in communityliving older adults: a 1-year prospective study. *Arch Phys Med Rehabil*. 2001;82(8):1050-1056.

5. Chase JAD, Phillips LJ, Brown M. Physical Activity Intervention Effects on Physical Function Among Community-Dwelling Older Adults: A Systematic Review and Meta-Analysis. *J Aging Phys Act*. 2017;25(1):149-170.

6. Evenson KR, Wen F, Herring AH. Associations of Accelerometry-Assessed and Self-Reported Physical Activity and Sedentary Behavior With All-Cause and Cardiovascular Mortality Among US Adults. *Am J Epidemiol*. 2016;184(9):621-632.

7. Bherer L, Erickson KI, Liu-Ambrose T. A review of the effects of physical activity and exercise on cognitive and brain functions in older adults. *J Aging Res*. 2013. doi: 10.1155/2013/657508.

8. Terwee CB, Bouwmeester W, van Elsland SL, de Vet HC, Dekker J. Instruments to assess physical activity in patients with osteoarthritis of the hip or knee: a systematic review of measurement properties. *Osteoarthr Cartil*. 2011;19(6): 620-633.

9. Nied RJ, Franklin B. Promoting and prescribing exercise for the elderly. *Am Fam Physician* 2002;65(3):419-426.

10. Svege I, Kollé E, Risberg MA. Reliability and validity of the Physical Activity Scale for the Elderly (PASE) in patients with hip osteoarthritis. *BMC Musculoskelet Disord*. 2012;13:26.

11. Forsén L, Loland NW, Vuillemin A, Chinapaw MJM, van Poppel, Mireille N M, et al. Self-administered physical activity questionnaires for the elderly: a systematic review of measurement properties. *Sports Med*. 2010;40(7):601-623.

12. Castañeda FJR, Aznar CT, Muro C, Guerra JC. Descripción de los instrumentos de medida de la movilidad en personas mayores de 65 años: revisión sistemática. *Rev. Esp. Salud Pública* 2015;89(6):545-561.

13. Geisinger KF. Cross-cultural normative assessment: Translation and adaptation issues influencing the nor-

- mative interpretation of assessment instruments. *Psychol Assess.* 1994;6(4):304-312.
14. Cukljek S, Juresa V, Babic J. The cross-cultural (trans-cultural) adaptation and validation of the nursing image questionnaire. *Nurse Educ Today.* 2017;48:67-71.
 15. Castellet E, Ares O, Celaya F, Valentí-Azcárate A, Salvador A, Torres A, et al. Transcultural adaptation and validation of the "Hip and Knee" questionnaire into Spanish. *Health Qual Life Outcomes.* 2014;12:76.
 16. Guirao-Goris JA, Cabrero-García J, Pina M, Patricia J, Muñoz-Mendoza CL. Revisión estructurada de los cuestionarios y escalas que miden la actividad física en los adultos mayores y ancianos. *Gac Sanit.* 2009;23(4):334.e51-334.e67.
 17. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. Protocol of the COSMIN study: Consensus-based Standards for the selection of health Measurement Instruments. *BMC Med Res Methodol.* 2006;6:2.
 18. Terwee CB, Mokkink LB, van Poppel MN, Chinapaw JM, van Mechelen W, de Vet HC. Qualitative attributes and measurement properties of physical activity questionnaires. *Sports Med.* 2010;40(7):525-537.
 19. World Health Organization. World report on ageing and health. Geneva: World Health Organization. 2015.
 20. World Health Organization. Global recommendations on physical activity for health. Geneva: World Health Organization. 2010.
 21. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health.* 1998;52(6):377-384.
 22. Prince SA, Adamo KB, Hamel ME, Hardt J, Connor Gorber S, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *Int J Behav Nutr Phys Act.* 2008;5:56.
 23. 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.
 24. Aggio D, Fairclough S, Knowles Z, Graves L. Validity and reliability of a modified english version of the physical activity questionnaire for adolescents. *Arch Public Health.* 2016;74:3.
 25. Anjana RM, Sudha V, Lakshmi Priya N, Subhashini S, Pradeepa R, Geetha L et al. Reliability and validity of a new physical activity questionnaire for India. *Int J Behav Nutr Phys Act.* 2015;12:40.
 26. De Abajo S, Larriba R, Marquez S. Validity and reliability of the Yale Physical Activity Survey in Spanish elderly. *J Sports Med Phys Fitness.* 2001;41(4):479-485.
 27. Donaire-Gonzalez D, Gimeno-Santos E, Serra I, Roca J, Balcells E, Rodríguez E, et al. Validación del cuestionario de actividad física de Yale en pacientes con enfermedad pulmonar obstructiva crónica. *Arch. Bronconeumol.* 2011;47(11):552-560.
 28. Dipietro L, Caspersen CJ, Ostfeld AM, Nadel ER. A survey for assessing physical activity among older adults. *Med Sci Sports Exerc.* 1993;25(5):628-642.
 29. Vilaró J, Gimeno E, Férez NS, Hernando C, Díaz I, Ferrer M, et al. Actividades de la vida diaria en pacientes con enfermedad pulmonar obstructiva crónica: validación de la traducción española y análisis comparativo de 2 cuestionarios. *Med Clin (Barc).* 2007;129(9):326-332.
 30. Voorrips LE, Ravelli AC, Dongelmans PC, Deurenberg P, Van Staveren WA. A physical activity questionnaire for the elderly. *Med Sci Sports Exerc.* 1991;23(8):974-979.
 31. Cancela JM, Ayán C, Sánchez-Lastra MA, Casal A. Construct Validity of the Spanish Version of the CHAMPS Questionnaire for institutionalized patients over 65 years old. *Motriz.* 2017;23(1):81-87.
 32. Stewart AL, Mills KM, Sepsis PG, King AC, McLellan BY, Roitz K, et al. Evaluation of CHAMPS, a physical activity promotion program for older adults. *Ann Behav Med.* 1997;19(4):353-361.
 33. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc.* 2003;35(8):1381-1395.
 34. Ruiz-Casado A, Alejo LB, Santos-Lozano A, Soria A, Ortega MJ, Pagola I, et al. Validity of the Physical Activity Questionnaires IPAQ-SF and GPAQ for Cancer Survivors: Insights from a Spanish Cohort. *Int J Sports Med.* 2016;37(12):979-985.
 35. Armstrong T, Bull F. Development of the World Health Organization Global Physical Activity Questionnaire (GPAQ). *J Public Health.* 2006;14(2): 66-70.
 36. Taylor HL, Jacobs DR, Schucker B, Knudsen J, Leon AS, Debacker G. A questionnaire for the assessment of leisure time physical activities. *J Chronic Dis.* 1978;31(12):741-755.
 37. Ruiz-Comellas A, Pera G, Baena-Díez JM, Mundet-Tudurí X, Alzamora-Sas T, Elosua R, et al. Validación de una versión reducida en español del cuestionario de

- actividad física en el tiempo libre de Minnesota (VREM). *Rev Esp Salud Publica*. 2012;86(5):495-508.
38. Cancela JM, Varela S, Alvarez MJ, Molina A, Ayán C, Martín V. Validity of a combined fibromyalgia (FM) questionnaires to assess physical activity levels in Spanish elderly women: an experimental approach. *Arch Gerontol Geriatr*. 2011;52(1):e56-9.
39. Mannerkorpi K, Hernelid C. Leisure Time Physical Activity Instrument and Physical Activity at Home and Work Instrument. Development, face validity, construct validity and test-retest reliability for subjects with fibromyalgia. *Disabil Rehabil*. 2005;27(12):695-701.
40. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420-428.
41. Nunnally JC, Bernstein IH. *Psychometric theory*, 3rd ed. New York: McGraw-Hill. 1994.
42. De Vet HC, Terwee CB, Knol DL, Bouter L. When to use agreement versus reliability measures. *J Clin Epidemiol*. 2006;59(10):1033-1039.
43. 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;13:32.
44. Pols MA, Peeters PH, Kemper HC, Grobbee D. Methodological aspects of physical activity assessment in epidemiological studies. *Eur J Epidemiol*. 1998;14(1):63-70.
45. 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.
46. Plasqui G, Westerterp KR. Physical activity assessment with accelerometers: an evaluation against doubly labeled water. *Obesity (Silver Spring)*. 2007;15(10):2371-2379.
47. Aggio D, Fairclough S, Knowles Z, Graves L. Validity and reliability of a modified english version of the physical activity questionnaire for adolescents. *Arch Public Health*. 2016;74:3.
48. 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(7):565-600.
49. Altman DG. *Practical statistics for medical research*. London: Chapman and Hall. 1991;589 p.
50. Harada ND, Chiu V, King AC, Stewart AL. An evaluation of three self-report physical activity instruments for older adults. *Med Sci Sports Exerc*. 2001;33(6):962-970.
51. Pettee Gabriel KK, Morrow JR, Woolsey AL. Framework for physical activity as a complex and multi-dimensional behavior. *J Phys Act Health*. 2012;9(Suppl 1):S11-S18.
52. Tudor-Locke CE, Myers AM. Challenges and opportunities for measuring physical activity in sedentary adults. *Sports Med*. 2001;31(2):91-100.