

# Validación de dos cuestionarios para evaluar el nivel de actividad física y el tiempo sedentario en una comunidad universitaria de Colombia

Validation of two questionnaires to assess the level of physical activity and sedentary time in a Colombian university community

Validação de dois questionários para avaliar o nível de atividade física e o tempo sedentário em uma comunidade universitária da Colômbia

Elkin Fernando Arango-Vélez<sup>1</sup>, Andrés Mauricio Echavarría-Rodríguez<sup>2</sup>, Fabián Alexander Aguilar-González<sup>3</sup>, Fredy Alonso Patiño-Villada<sup>4</sup>.

- 1 Epidemiólogo clínico, médico del Deporte, Grupo de Investigación en Actividad Física para la Salud (AFIS). Universidad de Antioquia, Colombia. elkinarango@yahoo.com. ORCID: <https://orcid.org/0000-0002-4017-9004>.
- 2 Licenciado en Educación Física, Grupo de Investigación en Actividad Física para la Salud (AFIS). Universidad de Antioquia, Colombia. brego91@gmail.com. ORCID: <https://orcid.org/0000-0002-0779-4806>.
- 3 Licenciado en Educación Física, Grupo de Investigación en Actividad Física para la Salud (AFIS). Universidad de Antioquia, Colombia. fabianalexagui@gmail.com. ORCID: <https://orcid.org/0000-0001-5907-5603>.
- 4 Doctor en Ciencias de la Actividad Física y del Deporte, licenciado en Educación Física, Grupo de Investigación en Actividad Física para la Salud (AFIS). Universidad de Antioquia, Colombia. fredy.patino@udea.edu.co. ORCID: <https://orcid.org/0000-0003-4419-0750>.

Recibido: 17/07/2018. Aprobado: 02/08/2019. Publicado: 05/02/2020

---

Arango-Vélez EF, Echavarría-Rodríguez AM, Aguilar-González FA, Patiño-Villada FA. Validación de dos cuestionarios para evaluar el nivel de actividad física y el tiempo sedentario en una comunidad universitaria de Colombia. Rev. Fac. Nac. Salud Pública. 2020;38(1):e334156. DOI: <https://doi.org/10.17533/udea.rfnsp.v38n1e334156>

---

## Resumen

**Objetivo:** Evaluar la validez concurrente y la confiabilidad test-retest del Global Physical Activity Questionnaire (GPAQ) y el International Physical Activity Questionnaire (IPAQ, versión corta) en una comunidad universitaria. **Metodología:** Un total de 76 personas de la Universidad de Antioquia, Medellín (Colombia), con edades entre 18 y 44 años, portaron un acelerómetro (ActiGraph GT3X+) durante siete días y posteriormente diligenciaron los dos cuestionarios; estos se repitieron una semana después. La validez y la confiabilidad para evaluar el gasto energético (GE) y el tiempo sedentario

(TS) se determinó con el coeficiente de correlación intraclase (CCI); la concordancia para clasificar el nivel de actividad física (NAF) se evaluó con el índice de Kappa ponderado (IKp). **Resultados:** Para medir el GE, el IPAQ mostró una validez “moderada” (CCI = 0,434, IC 95 % = 0,231-0,600;  $p < 0,001$ ) y el GPAQ “pobre” (CCI = 0,335, IC 95 % = 0,123-0,519;  $p < 0,001$ ); para el TS, la validez de ambos cuestionarios fue “pobre” (IPAQ: CCI = 0,234, IC 95 % = 0,011-0,435,  $p < 0,020$ ; GPAQ: CCI = 0,231, IC 95 % = 0,009-0,432,  $p < 0,021$ ). Para clasificar el NAF (bajo, moderado o alto),

ambos cuestionarios mostraron concordancias “bajas” (IPAQ:  $IK_p = 0,298$ ,  $p = 0,001$ ; GPAQ:  $IK_p = 0,137$ ,  $p = 0,119$ ). La confiabilidad del GPAQ para evaluar el GE fue “buena / excelente” (CCI = 0,838; IC 95 % = 0,754-0,895;  $p < 0,001$ ) y del IPAQ fue “moderada” (CCI = 0,655; IC 95 % = 0,505-0,766;  $p < 0,001$ ). La confiabilidad para el TS fue “moderada” en ambos instrumentos (IPAQ: CCI = 0,716; IC 95 % = 0,583-0,811;  $p < 0,001$ ; GPAQ: CCI = 0,736; IC 95 % = 0,613-0,824;

$p < 0,001$ ). **Conclusiones:** Los cuestionarios mostraron poca validez para la medición del NAF y del TS. Sin embargo, ambos instrumentos presentaron una buena confiabilidad para la medición de las variables estudiadas. No se recomienda el uso de los cuestionarios para la población estudiada.

-----**Palabras clave:** estudios de validación, actividad física, estilo de vida sedentario, Global Physical Activity Questionnaire, International Physical Activity Questionnaire.

## Abstract

**Objective:** Evaluate the concurrent validity and test-retest reliability of the Global Physical Activity Questionnaire (GPAQ) and the International Physical Activity Questionnaire (IPAQ, short version) in a university community. **Methodology:** A total of 76 people from the University of Antioquia, Medellín (Colombia), aged between 18 and 44, carried an accelerometer (ActiGraph GT3X+) for seven days and subsequently filled out the two questionnaires; these were repeated a week later. The validity and reliability for assessing energy expenditure (EE) and sedentary time (ST) was determined with the intraclass correlation coefficient (ICC); concordance to classify the physical activity level (PAL) was evaluated with the weighted Kappa index (wKI). **Results:** To measure the EE, IPAQ showed a “moderate” validity (ICC = 0.434, CI 95% = 0.231-0.600;  $p < 0.001$ ) and the “poor” GPAQ (ICC = 0.335, CI 95% = 0.123-0.519;  $p < 0.001$ ); for ST, the validity of both questionnaires was “poor” (IPAQ :

ICC = 0.234, CI 95% = 0.011-0.435,  $p < 0.020$ ; GPAQ: ICC = 0.231, CI 95% = 0.009-0.432,  $p < 0.021$ ). To classify the PAL (low, moderate or high), both questionnaires showed “low” concordance (IPAQ: wKI = 0.298,  $p = 0.001$ ; GPAQ: wKI = 0.137,  $p = 0.119$ ). The reliability of the GPAQ for assessing EE was “good/excellent” (ICC = 0.838; CI 95% = 0.754-0.895;  $p < 0.001$ ) and the IPAQ was “moderate” (ICC = 0.655; IC 95% = 0.505-0.766 ;  $p < 0.001$ ). The reliability for ST was “moderate” in both instruments (IPAQ: ICC = 0.716; CI 95% = 0.583-0.811;  $p < 0.001$ ; GPAQ: ICC = 0.736; CI 95% = 0.613-0.824;  $p < 0.001$ ). **Conclusions:** The questionnaires showed little validity for measuring PAL and ST. However, both instruments presented good reliability for measuring the variables studied. The use of questionnaires for the population studied is not recommended.

-----**Key words:** validation study, physical activity, sedentary lifestyle, Global Physical Activity Questionnaire, International Physical Activity Questionnaire

## Resumo

**Objetivo:** Avaliar a validade concorrente e a confiabilidade teste-reteste do Global Physical Activity Questionnaire (GPAQ) e do International Physical Activity Questionnaire (IPAQ, versão curta) em uma comunidade universitária. **Metodologia:** 76 pessoas da Universidade de Antioquia, Medellín (Colômbia), com idades entre 18 e 44 anos, carregaram um acelerômetro (ActiGraph GT3X+) por sete dias e, posteriormente, preencheram os dois questionários, que foram repetidos uma semana depois. A validade e a confiabilidade para avaliar o gasto de energia (GE) e o tempo sedentário (TS) foram determinadas com o coeficiente de correlação intraclasse (CCI); a concordância para classificar o nível de atividade física (NAF) foi avaliada com o índice Kappa ponderado (IK<sub>p</sub>). **Resultados:** Para medir o GE, o IPAQ mostrou uma validade “moderada” (CCI = 0,434, IC 95 % = 0,231-0,600;  $p < 0,001$ ) e o GPAQ “pobre” (CCI = 0,335, IC 95 % = 0,123-0,519;  $p < 0,001$ ); para o TS, a validade de ambos os questionários foi “pobre” (IPAQ: CCI = 0,234, IC 95 % =

0,011-0,435,  $p < 0,020$ ; GPAQ: CCI = 0,231, IC 95 % = 0,009-0,432,  $p < 0,021$ ). Para classificar o NAF (baixo, moderado ou alto), ambos os questionários mostraram concordâncias “baixas” (IPAQ:  $IK_p = 0,298$ ,  $p = 0,001$ ; GPAQ:  $IK_p = 0,137$ ,  $p = 0,119$ ). A confiabilidade do GPAQ para avaliar o GE foi “boa / excelente” (CCI = 0,838; IC 95 % = 0,754-0,895;  $p < 0,001$ ) e do IPAQ foi “moderada” (CCI = 0,655; IC 95 % = 0,505-0,766;  $p < 0,001$ ). A confiabilidade para o TS foi “moderada” em ambos os instrumentos (IPAQ: CCI = 0,716; IC 95 % = 0,583-0,811;  $p < 0,001$ ; GPAQ: CCI = 0,736; IC 95 % = 0,613-0,824;  $p < 0,001$ ). **Conclusões:** Os questionários mostraram pouca validade para a medição do NAF e do TS. No entanto, ambos os instrumentos apresentaram boa confiabilidade para a mensuração das variáveis estudadas. O uso de questionários para a população estudada não é recomendado.

-----**Palavras-chave:** estudos de validação, atividade física, estilo de vida sedentário, Global Physical Activity Questionnaire, International Physical Activity Questionnaire

## Introducción

La evidencia científica demuestra que los beneficios asociados a la práctica regular de actividad física (AF) se relacionan tanto con aspectos físicos como psicológicos; dichos beneficios se manifiestan en una disminución en el riesgo de padecer una creciente variedad de enfermedades crónicas, como la diabetes mellitus, los cánceres de colon y mama, la obesidad, la hipertensión arterial, la osteoporosis, la ansiedad, la depresión y el estrés, entre otras [1,2].

Sin embargo, datos mundiales indican que un 31,1 % de las personas de 15 años o más son físicamente inactivas, cifra que se acerca al 43 % en países americanos; además, la inactividad física aumenta con la edad, es más alta en las mujeres que en los hombres y también lo es en los países de ingresos económicos altos [3]. En Colombia, la prevalencia global de cumplimiento de las recomendaciones de AF entre las personas de 18 a 64 años es del 53,5 %, mientras que, en la AF del tiempo libre, solo 1 de cada 5 personas cumple con dichas recomendaciones [4]. En el ámbito universitario, datos nacionales muestran que un 22,0 % de los estudiantes reportan una AF insuficiente [5]; en un estudio realizado en la Universidad de Antioquia, la prevalencia de AF total baja fue del 28,7 % entre estudiantes, docentes y empleados [6].

Por otro lado, datos globales indican que cerca del 41,5 % de la población adulta pasa 4 o más horas del día sentados, datos que son muy similares en las Américas, con un 41,4 % [3]. El *tiempo sedentario* (TS) contempla todas aquellas actividades en las cuales las personas permanecen sentadas. Esto tiene una fuerte relación con un alto riesgo de mortalidad por enfermedad cardiovascular, diabetes tipo 2 y síndrome metabólico, riesgo que es independiente de la AF moderada o vigorosa que las personas reportan [7].

Para identificar tanto el *nivel de actividad física* (NAF) como el TS en las diferentes poblaciones, se requieren instrumentos que permitan medirlos. Los métodos objetivos (observación directa, calorimetría indirecta, agua doblemente marcada, acelerómetros, podómetros o la combinación de dispositivos) son herramientas válidas y confiables, pero no son de fácil acceso, debido a su costo y requieren de mucho tiempo para obtener la información de los sujetos [8,9]. De esta forma, se han desarrollado otros instrumentos de fácil acceso y bajo costo, como los cuestionarios de autorreporte, entre los cuales se encuentran el International Physical Activity Questionnaire (IPAQ), con sus dos versiones (corta y larga) [10], y el Global Physical Activity Questionnaire (GPAQ) [11]. Estas herramientas cuentan con estudios de validación y confiabilidad en diferentes contextos internacionales

[11-15]. Uno de estos estudios evaluó la validez y la confiabilidad del IPAQ en doce países, donde se sugiere que los resultados son comparables a nivel internacional; sin embargo, la validación de dicho cuestionario incluyó solo a Brasil como país suramericano [15]. Además, el uso de cuestionarios para evaluar el NAF y el TS suelen afectarse por factores externos, como son la convivencia social y las condiciones climáticas, las que pueden variar en cada contexto y deben considerarse para su aplicación [8].

En Colombia, estos cuestionarios se han empleado en diferentes poblaciones [4,6,16], pero no se tiene conocimiento de estudios que evaluaran la validez de estas herramientas a nivel nacional o local; solo un estudio reportó la reproducibilidad (prueba-reprueba) de ambos instrumentos en una población urbana adulta [17].

Una herramienta con buenos indicadores de validez y confiabilidad permite conocer, de forma práctica y precisa, el verdadero NAF y el TS de las personas, lo que a su vez posibilita diseñar programas más eficientes de AF y evaluar mejor su verdadero impacto. En el ámbito universitario, no se posee una herramienta validada para conocer los NAF y el TS, por lo que existe incertidumbre tanto sobre el impacto como por los beneficios de los programas orientados a la promoción de la AF.

Por lo anterior, este estudio tuvo como objetivo evaluar la validez concurrente y la confiabilidad test-retest de los cuestionarios GPAQ e IPAQ para medir el NAF y el TS, en un grupo de personas de una comunidad universitaria de Colombia.

## Metodología

### Diseño, población y muestra

Se efectuó un estudio cuantitativo transversal de validación de pruebas diagnósticas. La población de estudio fueron adultos jóvenes, con edades entre los 18 y 44 años, pertenecientes a la sede de Ciudadela Robledo, de la Universidad de Antioquia (Medellín, Colombia). Para el cálculo del tamaño de la muestra se consideró un coeficiente de correlación de 0,3 [11,15], una potencia del 80 % y un nivel de confianza del 95 %, lo que arrojó una muestra de 67 individuos, a la que se le sumó un 10 % de posibles pérdidas, para una muestra total de 74 personas. Se utilizó el software Epidat 3.1 para este cálculo. La muestra se tomó a conveniencia y se conformó con las personas que acudieron a la convocatoria abierta, que se llevó a cabo por medio de carteles, redes sociales y correos electrónicos.

El presente estudio cumplió con la normatividad nacional e internacional para investigación en humanos. El protocolo de investigación lo aprobó el Comité de Ética de la Investigación, del Instituto Universitario

de Educación Física de la Universidad de Antioquia (Resolución 113 de Consejo de Instituto del 1 de junio de 2015). Todos los participantes firmaron el consentimiento informado.

### Criterios de selección

Se incluyeron personas que tuvieran las siguientes características: algún vínculo con la Universidad de Antioquia (estudiante, docente o empleado), edad entre 18 y 44 años, y que aceptaran participar voluntariamente en el estudio. Se excluyó a quienes tuvieron al menos uno de los siguientes ítems: discapacidad física en miembros inferiores, practicar natación o deportes de contacto (por el riesgo de daño en los acelerómetros), procedimientos quirúrgicos en los últimos 3 meses, deportista de alto rendimiento y enfermedades crónicas descompensadas (hipertensión arterial, diabetes mellitus, asma, depresión, cáncer). También se excluyó a aquellas personas que mencionaban no contar con el tiempo para asistir a las evaluaciones requeridas durante la investigación.

### Control de sesgos

Los participantes fueron informados sobre los objetivos del estudio, la aplicación de los cuestionarios y el uso adecuado de los acelerómetros, para evitar la manipulación inadecuada de los mismos y la pérdida de datos. Los acelerómetros ActiGraph GT3X+ se verificaron antes de entregarlos a los participantes, con el fin de corroborar que se encontraran en estado óptimo para la recolección de los datos. Estos dispositivos se calibraron y programaron para recoger los datos durante un periodo de 7 días, el cual coincidió con el tiempo por el que indagaban los cuestionarios. Para el análisis de los resultados, se tuvo en cuenta el tiempo real de uso de los dispositivos.

Se realizó control de calidad de los cuestionarios que diligenciaron los participantes, para evitar la aparición de datos faltantes y respuestas erróneas; se definieron días específicos para la aplicación de los mismos, para minimizar los sesgos de memoria. Los datos obtenidos con los cuestionarios se analizaron de acuerdo con los parámetros establecidos en sus respectivas guías de análisis, y se aplicaron las sugerencias obtenidas en estudios previos sobre el manejo y uso de los mismos [10,18,19].

### Instrumentos y recolección de la información

Las variables sociodemográficas (sexo, edad, nivel socioeconómico, escolaridad, vínculo con la

Universidad) se recogieron por medio de una encuesta prediseñada. Las variables antropométricas se midieron así: peso, índice de masa corporal (IMC) y porcentaje de grasa, por medio de un dispositivo de bioimpedancia (Omron® HBF-510LA), cuya precisión es de 0,1 % para el porcentaje de grasa y de 100 gramos para el peso; la talla se evaluó con un estadiómetro de piso (precisión 0,1 cm), y el perímetro abdominal, con una cinta métrica.

Se utilizaron acelerómetros ActiGraph GT3X+ (Pensacola, Florida, Estados Unidos) los cuales se entregaron a los participantes con una carga superior a 4,1 voltios y su calibración permitió recoger los datos en los tres ejes del dispositivo (arriba abajo - derecha izquierda - adelante atrás), al sujetarse con una correa en el lado derecho de la cadera, sobre la línea axilar media.

Cada acelerómetro se programó con las siguientes características: la amplitud (capacidad del dispositivo para recoger datos) tuvo una frecuencia de 30 Hz; las épocas fueron de 60 segundos, con intervalos de tiempo en los que cada aceleración se registró en una señal digital cuantificable (*count*). Además, se ingresó la información biométrica que solicita el *software* (edad, peso, talla, sexo, raza, fecha de nacimiento, lado donde su ubicó el dispositivo y si este corresponde con el lado dominante).

La información de cada participante contenida en los acelerómetros se descargó y procesó con el *software* Actilife 6 (ActiGraph Pensacola, Florida, Estados Unidos). Se consideraron, como datos validados para ingresar en el análisis, las personas que presentaran un registro de actividad en los acelerómetros mayor o igual que 4 días, con un registro de *counts* mayor o igual que 10 horas por día [20]. Los análisis se realizaron con el vector de magnitud, que se obtuvo de la sumatoria de los datos en los tres ejes de medición. Para calcular el gasto energético (GE) en medida de unidad metabólica (*metabolic equivalent of task, MET*)\* / min / semana, se utilizó el algoritmo de Freedson para adultos [21], para cuyo cálculo se excluyó la actividad asociada a los períodos de no uso del acelerómetro y el tiempo de sueño, con el fin de tener datos que se pudieran comparar con los dominios de los cuestionarios GPAQ e IPAQ. Para el cálculo del TS se tomaron los periodos de 10 minutos consecutivos que oscilaran entre 0-99 counts / min; el TS obtenido se expresó en min / día.

Los cuestionarios con los cuales se cuantificó el NAF y el TS fueron el IPAQ para Colombia (versión corta telefónica) y el GPAQ versión 2 en español; ambos instrumentos fueron autoadministrados, con la asesoría

\* “Un MET es la relación entre la tasa metabólica de trabajo de una persona y su tasa metabólica de reposo. Un MET corresponde al costo energético de estar sentado tranquilamente y equivale a un consumo calórico de 1 kcal / kg / h. [...] comparado con estar tranquilamente sentado, el consumo calórico de una persona será cuatro veces más alto al realizar una actividad moderada y ocho veces mayor cuando realiza una actividad vigorosa” [18, p. 3; traducción nuestra].

de un investigador. La información que arrojaron los cuestionarios se analizó con las recomendaciones dadas en sus respectivas guías de análisis [10,18].

### Descripción de la evaluación

Luego de verificar el cumplimiento de los criterios de selección y obtener la firma del consentimiento informado, a cada participante se le realizaron las valoraciones de las variables sociodemográficas y antropométricas. Además, se les explicó, de forma detallada, el uso del acelerómetro, el cual portaron por 7 días, tiempo durante el cual a cada persona se le envió un mensaje por correo electrónico para recordarle usar el dispositivo. Luego de esto, diligenciaron los cuestionarios GPAQ e IPAQ, proceso que repitieron una semana después, con el fin de obtener los datos para calcular la confiabilidad de los mismos. La validez se evaluó con la primera medición de los cuestionarios.

### Análisis estadístico

Las variables cuantitativas se analizaron con la prueba de Kolmogorov Smirnov, para evaluar la distribución de las mismas; aquellas que tuvieron distribución normal se resumieron con medias y desviaciones estándar (DE), y las que no mostraron distribución normal, se resumieron con medianas y rangos intercuartílicos (RI).

Para analizar la validez y la confiabilidad (intrapueba) de cada uno de los cuestionarios, se utilizó el coeficiente de correlación intraclase (CCI); se tomó como una correlación “pobre” valores menores que 0,40; “moderada”, entre 0,41 y 0,75, y “buena” o “excelente”, mayores que 0,75 [22].

El GE semanal se estratificó así: NAF bajo: 0-600 MET / min / semana; NAF moderado: 601-1500 MET / min / semana; NAF alto: > 1500 MET / min / semana. Lo anterior se hizo para calcular la concordancia en la clasificación del NAF entre los acelerómetros y cada uno de los cuestionarios de AF, para lo que se utilizó el índice de Kappa ponderado (IKp), el cual se clasificó de la siguiente forma: “bajo”, menor que 0,40; “aceptable”, entre 0,41 y 0,60; “moderado”, entre 0,61 y 0,80; “excelente”, mayor que 0,80 [23].

Además, la concordancia se analizó con el método gráfico de Bland y Altman, el cual se construyó con las medias y las DE de las diferencias (acelerómetros - cuestionarios) entre las mediciones del GE y del TS de los cuestionarios IPAQ, GPAQ y los acelerómetros. Se obtuvieron los sesgos de medición y los límites de acuerdo (LA) [24].

Todos los análisis se hicieron con una significancia estadística de  $p < 0,05$  y una confiabilidad del 95 %. Se utilizó el paquete estadístico SPSS versión 24 (de International Business Machines Corporation, IBM®).

## Resultados

Los datos se recogieron entre agosto de 2015 y marzo de 2016. Inicialmente, 115 personas atendieron a la convocatoria y se les verificó los criterios de elegibilidad. De estas, fueron excluidas 37 personas: 8 que respondieron sí a la convocatoria inicial, pero que luego reusaron participar; 11 que eran deportistas de alto nivel o practicaban deportes donde no era recomendable usar el dispositivo, por el riesgo de ser dañado (fútbol, natación o deportes de contacto); 12 no contaban con el tiempo requerido para participar en el estudio (por prácticas externas o viajes a congresos), y 6 no tenían vínculo con la universidad o eran egresados.

Un total de 78 personas iniciaron su participación en el estudio, de los cuales dos fueron excluidos del análisis, por no cumplir con el tiempo de uso de los acelerómetros. De esta forma, se analizó la información obtenida de 76 personas, de quienes no se registró ningún dato perdido.

El promedio de uso de los acelerómetros de todos los participantes fue de 12,8 (DE = 1,5 horas / día). El 50 % de los participantes fueron mujeres, con una mediana de edad de 23 años (RI = 21,5-27,0); la mediana de educación formal fue de 16,5 años (RI = 15,0-18,3); 3 de cada 4 (75 %) de los sujetos mostró un IMC  $\leq 25,1$  kg / m<sup>2</sup>, y el 80,3 % fueron estudiantes (véase tabla 1).

El GE en los pretest, que se obtuvo con los cuestionarios y los acelerómetros, fue más bajo con el IPAQ (mediana = 1895,0 MET / minuto / semana; RI = 1016,0-3014,0), y el valor más alto lo arrojó el GPAQ (mediana = 2070,0 MET / minuto / semana; RI = 1320,0-3480,0). Por otro lado, el TS alcanzó medianas de 480 minutos / día, tanto con el GPAQ como con el IPAQ, mientras que en el re-test fueron de 420 minutos / día y 410 minutos / día, respectivamente (véase tabla 2). El NAF alto en los pretest fue el más prevalente tanto con los cuestionarios como con los acelerómetros, valor que fue levemente mayor en estos últimos, con un 73,3%.

### Validez concurrente

Al evaluar la validez de los cuestionarios, al compararlos con el “estándar de oro” (acelerómetros), se encontró que el IPAQ tiene una concordancia “moderada” (CCI = 0,434; IC 95% = 0,231-0,600;  $p < 0,001$ ), mientras para el GPAQ fue “pobre” (CCI = 0,335, IC 95% = 0,123-0,519;  $p < 0,001$ ) para cuantificar el GE. Para medir el TS, la concordancia de ambos cuestionarios fue “pobre” (IPAQ: CCI = 0,234, IC 95% = 0,011-0,435,  $p < 0,020$ ; GPAQ: CCI = 0,231, IC 95% = 0,009-0,432,  $p < 0,021$ ). Así mismo, para clasificar el NAF (bajo, moderado o alto), tanto el IPAQ (IKp = 0,298;  $p = 0,001$ ) como el GPAQ (IKp = 0,137;  $p = 0,119$ ) mostraron concordancias “bajas” al compararlos con los acelerómetros.

**Tabla 1.** Características sociodemográficas y antropométricas de los participantes (n = 76)

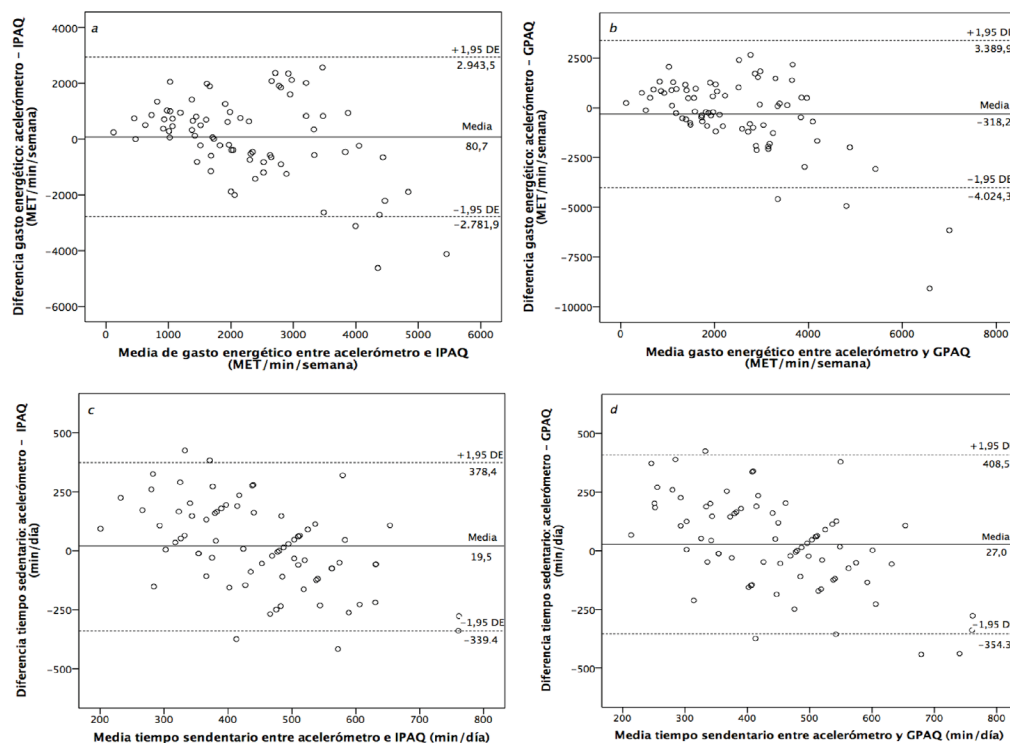
	Característica	Resultado	
		n	(%)
Sexo	Mujeres	38	50,0
	Hombres	38	50,0
Estrato socioeconómico	Bajo	28	36,8
	Medio	42	55,3
	Alto	6	7,9
Dependencia	Ciencias Agrarias	17	22,4
	Nutrición y Dietética	21	27,6
	Licenciatura en Educación Física	38	50,0
Vínculo	Empleado	5	6,6
	Docente	10	13,2
	Estudiante	61	80,3
	<b>Medianas</b>	<b>Rangos intercuartílicos (RI)</b>	
Edad (años)	23,0	(21,5-27,0)	
Talla (cm)	163,6	(158,8-174,0)	
Peso (kg)	62,6	(57,3-71,6)	
Índice de masa corporal (kg / m <sup>2</sup> )	23,4	(21,5-25,1)	
Perímetro abdominal (cm)	77,0	(73,0-82,8)	
Educación (años)	16,5	(15,0-18,3)	

**Tabla 2.** Descripción del gasto energético total, el tiempo sedentario y el nivel de actividad física (n = 76)

		Test		Re-test		
		Medianas	Rangos intercuartílicos (RI)	Medianas	Rangos intercuartílicos (RI)	
GE	IPAQ <sup>a</sup>	1895,0	(1016,0-3014,0)	1578,8	(1001,0-2679,0)	
GE	GPAQ <sup>a</sup>	2070,0	(1320,0-3480,0)	1900,0	(1160,0-3160,0)	
GE	acelerómetro <sup>a</sup>	2056,0	(1490,4-3203,3)		NA	
TS	IPAQ <sup>b</sup>	480,0	(300,0-600,0)	410,0	(240,0-540,0)	
TS	GPAQ <sup>b</sup>	480,0	(300,0-540,0)	420,0	(300,0-540,0)	
TS	acelerómetro <sup>b</sup>	472,5	(364,3-533,0)		NA	
		<b>n</b>	<b>(%)</b>	<b>n</b>	<b>(%)</b>	
NAF	IPAQ	Bajo	10	13,2	10	13,2
		Moderado	17	22,4	25	32,9
		Alto	49	64,5	41	53,9
NAF	GPAQ	Bajo	10	13,2	11	14,5
		Moderado	13	17,1	17	22,4
		Alto	53	69,7	48	63,2
NAF	Acelerómetro	Bajo	2	2,6		
		Moderado	18	23,7		NA
		Alto	56	73,7		

<sup>a</sup> MET / minuto / semana; <sup>b</sup> minutos / día.

GE = gasto energético; GPAQ = Global Physical Activity Questionnaire; IPAQ = International Physical Activity Questionnaire; NA = no aplica; NAF = nivel de actividad física; TS = Tiempo sedentario.



**Figura 1.** Gráficos de Bland-Altman para gasto energético (A y B) y tiempo sedentario (C y D), según datos de los acelerómetros y los cuestionarios IPAQ y GPAQ ( $n = 76$ ).

Los gráficos de Bland y Altman permiten ver el acuerdo entre los acelerómetros y los cuestionarios IPAQ corto y GPAQ para cuantificar el GE y el TS (véase figura 1). En el GE, la media de las diferencias entre el IPAQ y los acelerómetros mostró un sesgo positivo de 80,7 MET / min / semana (LA -2781,9 a 2.943,5), lo que indicó una tendencia a subestimar el GE con este cuestionario; mientras el GPAQ exhibió un sesgo negativo de -318,2 MET / min / semana (LA -4024,3 a 3.389,9), lo que señaló una sobrestimación del GE comparado con los acelerómetros. En el TS, ambos cuestionarios reportaron un sesgo positivo al compararlos con los acelerómetros; esto indicó que los cuestionarios tendieron a estimar en menor medida el TS (IPAQ: 19,5 minutos / día; LA -339,4 a 378,4; GPAQ: 27,0 minutos / día; LA -354,3 a 408,5). Los LA, tanto en los gráficos del GE como del TS, fueron amplios para ambos cuestionarios, pero mucho más con el GPAQ.

### Confiabilidad

En la confiabilidad intracuestionario (test-retest) para evaluar el GE, el GPAQ presentó una concordancia “buena / excelente” (CCI = 0,838; IC 95% = 0,754-0,895;  $p < 0,001$ ), clasificación que fue “moderada” para el IPAQ (CCI = 0,655; IC 95% = 0,505-0,766;

$p < 0,001$ ). Por otro lado, al calcular la confiabilidad intracuestionario para clasificar el NAF como alto, moderado o bajo, se halló una concordancia “aceptable” para el IPAQ (IKp = 0,484;  $p < 0,001$ ), mientras para el GPAQ fue “moderada” (IKp = 0,659;  $p < 0,001$ ). Respecto al TS, el análisis de confiabilidad intracuestionarios mostró una concordancia “moderada” para ambos instrumentos (IPAQ: CCI = 0,716; IC 95% = 0,583-0,811;  $p < 0,001$ ; GPAQ: CCI = 0,736; IC 95% = 0,613-0,824;  $p < 0,001$ ).

### Discusión

Los principales hallazgos de esta investigación fueron que, al comparar los cuestionarios y los acelerómetros, la validez concurrente para evaluar el GE del IPAQ corto alcanzó a ser “moderada”, mientras la del GPAQ fue “pobre”. Por otro lado, la validez de ambos cuestionarios para cuantificar el TS fue “pobre”, y para clasificar el NAF fue “baja”. La confiabilidad en la cuantificación del GE durante la AF diaria fue “buena / excelente” para el GPAQ y “moderada” para el IPAQ corto, mientras que para la clasificación del NAF, el acuerdo fue “moderado” y “aceptable”, respectivamente. Por último,

la confiabilidad para evaluar el TS fue “moderada” para ambos cuestionarios.

### Validez de criterio

Una característica importante de cualquier herramienta diagnóstica es la validez de esta para cuantificar la variable que se pretende medir al compararla con un “estándar de oro”, lo que se conoce como “validez de criterio”. En este estudio se encontró una concordancia “moderada” para el IPAQ corto y “pobre” con el GPAQ, en la evaluación del GE, al compararlos con los acelerómetros.

Con respecto al GPAQ, estos hallazgos se asemejan con los reportados en un grupo de estudiantes hombres universitarios de Arabia Saudita [25] y en una investigación de validación que se realizó en nueve países [11], en los cuales la validez de dicho instrumento fue “pobre”; y difieren de un estudio que incluyó personas de ambos sexos con un promedio de edad de 44 años, cuya validez fue “moderada” ( $\rho = 0,484$ ;  $p < 0,005$ ), al compararlo con acelerómetros [26].

Así mismo, tanto en una revisión sistemática que se publicó en 2011, donde hallaron una mediana para el coeficiente de correlación de Spearman (CCS) de 0,28 (RI = 0,09-0,39) para el IPAQ corto, como en un estudio poblacional en Noruega, donde dicha correlación fue 0,33, los resultados fueron consistentes con los que se obtuvieron en esta investigación [27,28]. Se debe tener en cuenta que, en dichas investigaciones, el estadístico que utilizaron para reportar la validez fue el CCS y tomaron como medida de comparación los minutos / semana de AF [11,25,27].

En lo referente a clasificar el NAF, tanto el IPAQ corto como el GPAQ arrojaron una validez “baja”, al ser comparados contra los acelerómetros, característica que puede llevar a errores importantes cuando se utilicen estas herramientas para evaluar el cumplimiento de las recomendaciones internacionales de AF, situación que se corrobora con el reporte de un estudio de validación del IPAQ corto en Cataluña (España), en el cual, para esta indicación, fue “bajo” ( $k = 0,33$ ,  $p < 0,05$ ) [29].

Respecto a la cuantificación del TS, ambos cuestionarios presentaron una validez “pobre” con respecto a los acelerómetros. Estos resultados coincidieron con los obtenidos en otros estudios [25,26,30]. Sin embargo, una investigación que empleó el IPAQ e involucró a 1751 adultos (19-84 años), reportó una validez moderada (CCS = 0,46;  $p > 0,0001$ ) [28]. En la actualidad, la medición del TS es tan importante como la cuantificación de la AF, debido a que el aumento del comportamiento sedentario aumenta el riesgo de mortalidad por enfermedades crónicas, en la misma medida que un bajo NAF [7].

Los anteriores hallazgos son consistentes con la evidencia que se obtuvo al construir los gráficos de

Bland y Altman, en los cuales, para medir el GE, el IPAQ corto mostró un sesgo positivo promedio de 80,7 MET / min / semana, y para el GPAQ, dicho sesgo fue negativo (-318,2 MET / min / semana); ambos cuestionarios, con unos límites de acuerdo muy amplios. Un hallazgo semejante se obtuvo con estos gráficos para evaluar el TS, los cuales mostraron sesgos positivos (IPAQ corto = 19,5 min / día; GPAQ = 27,0 min / día), con límites de acuerdo muy amplios. Dichos resultados permiten ver que estas herramientas arrojan valores en las mediciones que no concuerdan con las que brindan los acelerómetros, situación que dificulta la interpretación y la toma de decisiones con los datos tanto del IPAQ corto como del GPAQ. En la literatura se encuentran investigaciones que dan resultados semejantes a este estudio, tanto para la práctica de AF como para el TS, lo que se evidencia en que los límites de acuerdo entre los acelerómetros y los cuestionarios son bastante amplios, a pesar de que en dichas publicaciones las comparaciones entre los instrumentos se realizaron con los minutos / día de AF moderada-vigorosa [25,27,31].

### Confiabilidad

Para evaluar el GE, el GPAQ mostró una confiabilidad test-retest “buena / excelente”, resultado casi idéntico al que reportaron Chu et al. con un CCI de 0,79 (IC 95% 0,75 a 0,91) [31]; así mismo, para el IPAQ corto, la confiabilidad test-retest alcanzó a ser “moderada”, resultado semejante al que reportaron Silsbury, Goldsmith y Rushton en una revisión sistemática de 2015 [32] y en un estudio multicéntrico que se publicó en 2003 [15].

El GPAQ mostró una confiabilidad “moderada” para clasificar el NAF, la cual fue superior a la que se reportó en un estudio ( $n = 56$ ; mediana edad 32 años; RI 23-67), en el que alcanzó a ser “aceptable” ( $K_p = 0,41$ ; IC 95% = 0,21-0,61) [31]. Así mismo, el IPAQ corto arrojó una confiabilidad “aceptable” para dicha clasificación.

Respecto a la evaluación del TS, la confiabilidad alcanzada por ambos instrumentos fue “moderada”. Estos resultados están acordes con los reportados en un estudio realizado en Bucaramanga (Colombia) con personas de 18 a 70 años, que presentó una confiabilidad “moderada” para el IPAQ (CCI = 0,77 IC 95% = 0,67-0,84) y “excelente” para el GPAQ (CCI: 0,83 IC 95% = 0,76-0,89) [17]. Igualmente, en el estudio de Alkahtani [25], se informó una confiabilidad “moderada” para el GPAQ, para la cuantificación del TS.

Las diferencias de los resultados de validez y confiabilidad de este estudio con los reportes de otras investigaciones pueden explicarse por los diversos dispositivos que se utilizaron en ellas, es decir, cada una empleó un acelerómetro de características diferentes; además, pudieron intervenir la cultura y los niveles de educación de los participantes [27,28]. A pesar de su



escaza validez, los cuestionarios para medir los NAF y el TS tienen como ventajas el ser de bajo costo, fáciles de aplicar y el no alterar el comportamiento de las personas a quienes se les aplica [33].

Lo que se puede recomendar cuando se emplean estos cuestionarios, pese a su moderada confiabilidad, es que se debe reconocer un potencial sesgo de información, ya que al utilizar estos instrumentos en estudios poblacionales, se puede estar brindando estimaciones de la prevalencia de AF muy alejadas de la verdad, que pueden tener implicaciones en la toma de decisiones de la salud pública. Lo que debe procurarse hacer, a pesar del mayor costo, es la utilización de medidas objetivas para la cuantificación de la AF en la evaluación de programas, que promueven estilos de vida activos como los desarrollados en el ámbito universitario. Esto está soportado en Silfee et al. [34], quienes indican un aumento en la utilización de medidas objetivas de un 4,4 al 70,6 % entre 2006 y 2016 para la evaluación de intervenciones en AF; y, además, señalan la necesidad de seguir incorporando estas medidas en la investigación de la AF.

### Fortalezas y limitaciones

Este estudio tiene, como fortalezas, el uso de pruebas estadísticas robustas para determinar el grado de concordancia entre los cuestionarios y el “estándar de oro”, como son el CCI y el índice de Kappa ponderado, mientras que en la mayoría de los estudios de validación del IPAQ y el GPAQ publicados usaron pruebas estadísticas, como los coeficientes de correlación de Pearson y de Spearman, y las diferencias de medias con la prueba t de student, las cuales no evalúan concordancia o nivel de acuerdo, que es, en últimas, la característica clave para evaluar la validez de una prueba diagnóstica [22].

Entre las limitaciones de este estudio se debe tener presente que el “estándar de oro” empleado (acelerómetros) tiene una validez baja para medir la AF. Estos dispositivos subestiman algunas actividades físicas que se realizan durante la vida diaria, como son los trabajos en el hogar y el transporte en bicicleta, las que llevan a poco movimiento del centro de gravedad y, por ende, no son detectadas por el dispositivo [9]. Sin embargo, un estudio previo indicó como apropiado el uso del acelerómetro ActiGraph GT3X+, por considerarse un modelo que demuestra confiabilidad tanto en el ámbito clínico como en estudios epidemiológicos, además de ser el único modelo comercial que reporta correlaciones significativas con la técnica de agua doblemente marcada en la medición del gasto energético [26].

La aplicación de cuestionarios autoadministrados en comparación con los empleados por medio de entrevista podría considerarse como otra limitación. Frente a

esto, un estudio sobre el IPAQ, que comparó ambas formas de administración del instrumento, indicó que la versión autoadministrada presentó resultados superiores en el reporte de la AF y menores en el TS comparado con la entrevista [35]; sin embargo, estas diferencias fueron mayores en hombres, en personas con bajo nivel educativo, adultos mayores y en personas con sobrepeso u obesidad, población con características diferentes a la del presente estudio, con excepción de los hombres. Otro estudio sobre el IPAQ concluyó que se obtienen puntuaciones más altas con el autoinforme, posiblemente debido a la mala interpretación del instrumento en este modo de empleo [36]; sin embargo, en el actual estudio, las personas no estuvieron totalmente solas, ya que fueron asesoradas por un investigador en el momento de diligenciar el instrumento. Respecto al GPAQ, un estudio indicó que ambas formas de empleo son intercambiables, con una ventaja en la disminución de costos por parte del cuestionario autoadministrado [31], algo que era necesario en esta investigación por no contar con financiación.

Otra posible limitación es que nuestro estudio empleó el gasto energético como medida de comparación entre los cuestionarios y los acelerómetros, mientras los demás estudios encontrados en la literatura utilizaron la medición de la AF de intensidad moderada a vigorosa, expresada en min / semana; sin embargo, estos estudios igualmente reportaron una baja validez por parte de los cuestionarios [25,27,28], con lo que es posible afirmar que los resultados de validez son independientes de la unidad de medida empleada para evaluar la concordancia de estos instrumentos.

Por otro lado, el rango de edad con que se trabajó en esta investigación limita la validez externa de los resultados, los cuales solo podrán aplicarse a personas con características semejantes a los participantes que se incluyeron.

Finalmente, una de las limitaciones del estudio fue el haber trabajado con una muestra a conveniencia, no aleatoria, lo que puede llevar a sesgos de selección de los participantes.

## Conclusiones

Los cuestionarios IPAQ corto y GPAQ tienen una validez de criterio “pobre” para medir el GE total, el TS y clasificar el NAF en una comunidad universitaria de la ciudad de Medellín. Sin embargo, ambos instrumentos presentan una buena confiabilidad para la medición de las variables estudiadas.

Aunque se conoce la utilidad práctica de los cuestionarios, su aplicación se ve limitada por la forma en que se aleja de la verdad que pretender medir. Para la población estudiada se recomienda no hacer uso de

estos instrumentos o buscar otras herramientas para cuantificar la AF y el TS.

## Conflicto de interés

Los investigadores declaran no tener ningún conflicto de interés.

## Financiación

La presente investigación no contó con financiación. La Universidad de Antioquia designó tiempo a los profesores investigadores, y el Laboratorio Integrado de Ciencias Aplicadas a la Actividad Física y el Deporte, de la misma universidad, facilitó el préstamo de los acelerómetros para la toma de los datos.

## Declaración de responsabilidad

Se declara que los puntos de vista expresados son responsabilidad de los autores y no de la institución en la que trabajan.

## Referencias

- Garber CE, Blissmer B, Deschenes MR, et al. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: Guidance for prescribing exercise. *Med Sci Sport Exerc.* 2011;43(7):1334-59. DOI: <https://doi.org/10.1249/MSS.0b013e318213fefb>.
- Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: The evidence. *CMAJ.* 2006;174(6):801-9. DOI: <https://doi.org/10.1503/cmaj.051351>.
- Hallal PC, Andersen LB, Bull FC, et al. Global physical activity levels: Surveillance progress, pitfalls, and prospects. *Lancet.* 2012;380(9838):247-57. DOI: [https://doi.org/10.1016/S0140-6736\(12\)60646-1](https://doi.org/10.1016/S0140-6736(12)60646-1).
- González S, Lozano O, Ramírez A, et al. Niveles de actividad física de la población colombiana: desigualdades por sexo y condición socioeconómica. *Biomédica.* 2014;34(3):447-59. DOI: <http://dx.doi.org/10.7705/biomedica.v34i3.2258>.
- Varela MT, Duarte C, Salazar IC, et al. Actividad física y sedentarismo en jóvenes universitarios de Colombia: prácticas, motivos y recursos para realizarlas. *Colomb Med.* 2011;42(3):269-77.
- Arboleda-Serna V, Arango-Vélez E, Feito Y. Niveles y estados de cambio de la actividad física en una comunidad Universitaria de Medellín-Colombia. *Educ Fis Deport.* 2014;33(1):153-73. DOI: <http://doi.org/10.17533/udea.efyd.v33n1a09>.
- De Rezende LFM, Rodrigues Lopes M, Rey-Lopez JP, et al. Sedentary behavior and health outcomes: An overview of systematic reviews. *PLoS One.* 2014;9(8):e105620. DOI: <http://doi.org/10.1371/journal.pone.0105620>.
- Sylvia LG, Bernstein EE, Hubbard JL, et al. Practical guide to measuring physical activity. *J Acad Nutr Diet.* 2014;114(2):199-208. DOI: <http://doi.org/10.1016/j.jand.2013.09.018>.
- Ainsworth B, Cahalin L, Buman M, et al. The current state of physical activity assessment tools. *Prog Cardiovasc Dis.* 2015;57(4):387-95. DOI: <http://doi.org/10.1016/j.pcad.2014.10.005>.
- International Physical Activity Questionnaire (IPAQ). Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire [internet]. 2005 [citado 2015 ago. 15]. Disponible en: <https://sites.google.com/site/theipaq/scoring-protocol>.
- Bull FC, Maslin TS, Armstrong T. Global physical activity questionnaire (GPAQ): Nine country reliability and validity study. *J Phys Act Heal.* 2009;6(6):790-804.
- Trinh OT, Nguyen ND, Van der Ploeg HP, et al. Test-retest repeatability and relative validity of the Global Physical Activity Questionnaire in a developing country context. *J Phys Act Heal.* 2009;6 Suppl 1:S46-53.
- Medina C, Barquera S, Janssen I. Validity and reliability of the International Physical Activity Questionnaire among adults in Mexico. *Rev Panam Salud Pública.* 2013;34(1):21-28.
- Chun MY. Validity and reliability of korean version of international physical activity questionnaire short form in the elderly. *Korean J Fam Med.* 2012;33(3):144-51. DOI: <http://doi.org/10.4082/kjfm.2012.33.3.144>.
- Craig CL, Marshall AL, Sjoström M, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sport Exerc.* 2003;35(8):1381-95. DOI: <http://doi.org/10.1249/01.MSS.0000078924.61453.FB>.
- Sarmiento OL, Schmid TL, Parra DC, et al. Quality of life, physical activity, and built environment characteristics among colombian adults. *J Phys Act Health.* 2010;7 Suppl 2(Suppl 2):S181-95.
- Angarita A, Camargo D, Oróstegui M. Reproducibilidad del tiempo en posición sedente evaluado con el International Physical Activity Questionnaire (IPAQ) y el Global Physical Activity Questionnaire (GPAQ). *MedUNAB.* 2010;3(1):5-12.
- WHO. Global physical activity questionnaire (GPAQ). Analysis Guide [internet]. 2013 [citado 2015 ago. 15]. Disponible en: [https://www.who.int/ncds/surveillance/steps/resources/GPAQ\\_Analysis\\_Guide.pdf](https://www.who.int/ncds/surveillance/steps/resources/GPAQ_Analysis_Guide.pdf).
- Hallal PC, Gomez LF, Parra DC, et al. Lessons learned after 10 years of IPAQ use in Brazil and Colombia. *J Phys Act Heal.* 2010;7 Suppl 2:S259-64.
- Migueles JH, Cadenas-Sanchez C, Ekelund U, et al. Accelerometer data collection and processing criteria to assess physical activity and other outcomes: A systematic review and practical considerations. *Sport Med.* 2017;47(9):1821-45. DOI: <http://doi.org/10.1007/s40279-017-0716-0>.
- Lyden K, Kozey SL, Staudenmeyer JW, et al. A comprehensive evaluation of commonly used accelerometer energy expenditure and MET prediction equations. *Eur J Appl Physiol.* 2011;111(2):187-201. DOI: <http://doi.org/10.1007/s00421-010-1639-8>.
- Martínez-González M, Sánchez-Villegas A, Toledo-Atucha E, et al. *Bioestadística amigable.* 3.ª ed. Barcelona: Elsevier; 2014.
- Streiner DL, Norman GR, Cairney J. *Health Measurement Scales: A Practical Guide to their Development and Use.* 5.ª ed. Oxford: Oxford University Press; 2015.
- Giavarina D. Understanding Bland Altman analysis. *Biochem Med.* 2015;25(2):141-51. DOI: <http://doi.org/10.11613/BM.2015.015>.

25. Alkahtani SA. Convergent validity: Agreement between accelerometry and the Global Physical Activity Questionnaire in college-age Saudi men. *BMC Res Notes*. 2016;9(1):436. DOI: <http://doi.org/10.1186/s13104-016-2242-9>.
26. Cleland CL, Hunter RF, Kee F, et al. Validity of the Global Physical Activity Questionnaire (GPAQ) in assessing levels and change in moderate-vigorous physical activity and sedentary behaviour. *BMC Public Health*. 2014;14(1). DOI: <http://doi.org/10.1186/1471-2458-14-1255>.
27. Lee PH, Macfarlane DJ, Lam TH, et al. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): A systematic review. *Int J Behav Nutr Phys Act*. 2011;8(1):115. DOI: <http://doi.org/10.1186/1479-5868-8-115>.
28. Dyrstad SM, Hansen BH, Holme IM, et al. Comparison of self-reported versus accelerometer-measured physical activity. *Med Sci Sports Exerc*. 2014;46(1):99-106. DOI: <http://doi.org/10.1249/MSS.0b013e3182a0595f>.
29. Román Viñas B, Ribas Barba L, Ngo J, et al. Validación en población catalana del cuestionario internacional de actividad física. *Gac Sanit*. 2013;27(3):254-7. DOI: <http://dx.doi.org/10.1016/j.gaceta.2012.05.013>.
30. Curry WB, Thompson JL. Comparability of accelerometer- and IPAQ-derived physical activity and sedentary time in South Asian women: A cross-sectional study. *Eur J Sport Sci*. 2015;15(7):655-62. DOI: <http://doi.org/10.1080/17461391.2014.957728>.
31. Chu AHY, Ng SHX, Koh D, et al. Reliability and validity of the self- and interviewer-administered versions of the Global Physical Activity Questionnaire (GPAQ). *PLoS One*. 2015;10(9):e0136944. DOI: <https://doi.org/10.1371/journal.pone.0136944>.
32. Silsbury Z, Goldsmith R, Rushton A. Systematic review of the measurement properties of self-report physical activity questionnaires in healthy adult populations. *BMJ Open*. 2015;5(9):e008430. DOI: <https://doi.org/10.1136/bmjopen-2015-008430>.
33. Terwee CB, Mokkink LB, Van Poppel MN, et al. Qualitative attributes and measurement properties of physical activity questionnaires: A checklist. *Sport Med*. 2010;40(7):525-37. DOI: <https://doi.org/10.2165/11531370-000000000-00000>.
34. Silfee VJ, Haughton CF, Jake-Schoffman DE, et al. Objective measurement of physical activity outcomes in lifestyle interventions among adults: A systematic review. *Prev Med Rep*. 2018;11:74-80. DOI: <https://doi.org/10.1016/j.pmedr.2018.05.003>.
35. Van Dyck D, Cardon G, Deforche B, et al. IPAQ interview version: Convergent validity with accelerometers and comparison of physical activity and sedentary time levels with the self-administered version. *J Sports Med Phys Fitness*. 2015;55(7-8):776-86.
36. Bandeira F de M, Freitas MP, László M, et al. Mode of administration does matter: comparability study using IPAQ. *Motriz: Rev Educ Física*. 2015;21(4):370-4. DOI: <https://doi.org/10.1590/S1980-65742015000400005>.