

Validez de criterio de las escalas de medición de esfuerzo percibido en niños sanos: una revisión sistemática y metaanálisis

Criterion-related validity of perceived exertion scales in healthy children: a systematic review and meta-analysis

- a. Escuela de Kinesiología, Facultad de Ciencias de la Salud, Universidad San Sebastián, Concepción, Chile.
- b. Carrera de Kinesiología, Facultad de Salud, Universidad Santo Tomás, Concepción, Chile.
- c. Programa de Doctorado en Ciencias Médicas, Universidad de la Frontera, Temuco, Chile.
- d. Programa de Doutorado em Farmacologia, Departamento de Fisiologia, Escola Paulista de Medicina, Universidade Federal de São Paulo, Brasil.
- e. Departamento de Morfología, Facultad de Ciencias Médicas, Universidad Nacional Autónoma de Honduras (UNAH), Honduras.
- f. Departamento de Cirugía, Universidad de La Frontera, Chile.
- g. Centro de Investigación en Ciencias Biomédicas, Universidad Autónoma de Chile.

Klgo. Iván Rodríguez^{a,b,c,d}, Dra. Lysien Zambrano^{d,e} y Dr. Carlos Manterola^{c,f,g}

RESUMEN

Introducción. Los parámetros fisiológicos para medir la intensidad de ejercicio son el consumo de oxígeno y la frecuencia cardíaca. Sin embargo, la medición del esfuerzo percibido (EP) también es, a menudo, utilizada. El objetivo del estudio es determinar la validez asociada a criterio de las escalas de EP en niños durante una prueba de ejercicio incremental.

Métodos. Siete bases de datos electrónicas fueron utilizadas. Se incluyeron estudios destinados a evaluar la validez de criterio de las escalas de EP en niños sanos durante un test de ejercicio incremental. Los coeficientes de correlación fueron transformados a valores z y evaluados por un metaanálisis mediante un modelo de efectos fijos o efectos aleatorios, dependiendo de si el I^2 fue menor o mayor de 50%, respectivamente.

Resultados. Veinticinco artículos, que consideraron 1418 niños (49,2% varones), cumplieron los criterios de inclusión. El promedio de edad fue 10,5 años. Los modos de ejercicio observados fueron bicicleta, carrera y ejercicios de escalón. El coeficiente de correlación ponderado fue de 0,835 (IC 95%: 0,762-0,887) y 0,874 (IC 95%: 0,794-0,924) considerando frecuencia cardíaca y consumo de oxígeno como criterio de referencia. El paradigma de producción y aquellas escalas sin adaptaciones para niños mostraron el menor rendimiento de medición ($p < 0,05$).

Conclusión. La medición de EP sería válida en niños sanos durante un test de ejercicio incremental. Las escalas de medición específicas para niños mostraron mejor rendimiento que aquellas escalas no adaptadas. Estudios con mejor calidad metodológica deben ser realizados para confirmar estos resultados.

Palabras clave: validez, niños, escalas, cognición, ejercicio.

<http://dx.doi.org/10.5546/aap.2016.120>

INTRODUCCIÓN

El ejercicio físico es importante para mantener una buena salud. Un apropiado método de evaluación de su intensidad contribuye a lograr los umbrales fisiológicos que permiten estimular los mecanismos de adaptación inducidos por el

ejercicio.¹ Las variables fisiológicas generalmente consideradas en la evaluación y prescripción del ejercicio son consumo de oxígeno (VO_2), frecuencia cardíaca (FC), lactato sanguíneo y frecuencia respiratoria (FR).¹ No obstante, dado que la medición de estas variables requiere de un equipamiento costoso y de la supervisión de expertos, la percepción de esfuerzo (PE) se utiliza con frecuencia.²

La PE puede ser considerada como una configuración de sensaciones que integra tensión, dolor y fatiga, que vincula los sistemas muscular, cardiovascular y pulmonar durante el ejercicio.² A medida que la intensidad del ejercicio aumenta, existe un correspondiente e interdependiente incremento en la magnitud de la respuesta a lo largo de un continuum perceptual y fisiológico, lo que demuestra la existencia de una fuerte correlación positiva.² Esta asociación ha permitido la utilización del EP como variable, tanto para estimar la carga de trabajo como para producir un determinado nivel de intensidad durante el ejercicio.³ Por esta razón, constituye una herramienta de gran utilidad clínica, puesto que, mediante un método sencillo y costo-efectivo, es posible determinar y dosificar la intensidad de ejercicio con un alto nivel de certeza.^{2,3}

La escala de Borg es una de las escalas de EP más frecuentemente utilizadas y ha mostrado ser una herramienta válida en sujetos adultos.³ Sin embargo, en niños, esta escala no ha probado tener propiedades psicométricas favorables.⁴⁻⁷ En

Correspondencia:
Klgo. Iván Rodríguez,
ivan.rodriguez@uss.cl

Financiamiento:
Ninguno

Conflicto de intereses:
Ninguno que declarar.

Recibido: 30-6-2015
Aceptado: 7-9-2015

virtud de lo anterior, varias escalas lineales han sido desarrolladas para niños sobre la base de expresiones comunes y con un número limitado de rangos (0-10). Estas han sido utilizadas con diversos grados de éxito en la medición del EP.² Aparentemente, han mostrado propiedades psicométricas adecuadas en niños.

De acuerdo con nuestro conocimiento, no existen estudios dirigidos a determinar la validez de estas escalas en niños mediante un enfoque metaanalítico. En virtud de lo anterior, el objetivo de este estudio es determinar la validez asociada a criterio de las escalas de EP en niños, así como también evaluar el impacto de algunas covariables asociadas a esta propiedad de medición.

MÉTODOS

El presente estudio corresponde a una revisión sistemática (RS).⁸

Criterios de elegibilidad de los artículos: Se consideraron estudios realizados en humanos con diseño correlacional publicados en idioma inglés, portugués, francés o español. Los participantes fueron definidos como sujetos sanos menores de 18 años de ambos sexos. Se excluyeron aquellos estudios que agruparon los datos de adultos o niños con alguna patología concomitante. Los estudios tenían que ser enfocados en determinar la correlación entre el EP y las variables fisiológicas (criterios de referencia) durante un test de ejercicio incremental (TEI). Como criterios de referencia, se consideraron FC, VO_2 , carga de trabajo, FR, ventilación minuto (VE), equivalente ventilatorio para el oxígeno (VE/VO_2) y cociente respiratorio (VCO_2/VO_2).

Búsqueda de artículos: En las bases de datos Pubmed, ProQuest, *Scientific Electronic Library Online* (SciELO), SPORTDiscus, *Rehabilitation and Sports Medicine Source* (R&SMS), *Cumulative Index to Nursing & Allied Health Literature* (CINAHL) y Trip data base, se consideraron artículos desde su ingreso hasta abril de 2015. Los términos descriptores médicos (*Medical Subject Headings*; MeSh, por sus siglas en inglés) (niños, adolescentes, ejercicio, test de ejercicio, disnea, cognición, frecuencia cardíaca, consumo de oxígeno, volumen corriente, ventilación pulmonar y frecuencia respiratoria) y términos libres (percepción de esfuerzo, intensidad de ejercicio, validez, validez concurrente) fueron utilizados. También se utilizaron los operadores booleanos AND y OR. Adicionalmente, se revisaron las listas de referencia de los artículos primarios. La búsqueda fue realizada entre febrero y abril de 2015.

Identificación de los artículos: Se revisaron los títulos y los resúmenes de los artículos atingentes. Luego se obtuvieron los textos extensos de los resúmenes que cumplieron con los criterios de elegibilidad y los artículos atingentes de las listas de referencia. Los datos fueron enmascaradamente extraídos por dos investigadores independientes (IR, LZ) y registrados en una planilla *ad hoc*. Se registró año y lenguaje de publicación, tamaño de muestra, edad de los participantes, test de ejercicio realizado, escala de EP evaluada, coeficiente de correlación (CC) y criterio de referencia.

Evaluación de la calidad metodológica y riesgo de sesgo: La calidad metodológica (CM) fue evaluada de manera independiente por dos revisores (IR, LZ) a través de la escala *Consensus-based Standards for the Selection of Health Measurement* (COSMIN). La escala COSMIN permite determinar la CM de los estudios dirigidos a evaluar las propiedades psicométricas de los parámetros de medición en salud. Para este estudio, solo fue considerada la sección de la escala enfocada en evaluar la CM de estudios de validez de criterio.⁹

El riesgo de sesgo fue evaluado mediante la correlación entre el CC y el tamaño muestral. Para esto, se construyó un gráfico de embudo^{3,10} (*Anexo en formato electrónico*) y se calculó el CC de rangos de Kendall (prueba de correlación de rangos de Begg y Mazumdar).

Análisis estadístico. Cuando el VO_2 fue reportado en valores absolutos (ml/min) y relativos (ml/kg/min) en el mismo artículo, se calculó el promedio ponderado del CC (de acuerdo con el tamaño muestral). Adicionalmente, cuando el resultado de la asociación fue presentado como coeficiente de determinación (R^2), se calculó la raíz cuadrada con el fin de obtener el CC r de Pearson.

El metaanálisis fue realizado de acuerdo con el método de Hedges-Olkin, en el que se realiza la transformación z de Fisher de los CC.³ Por su parte, la inconsistencia fue calculada a través del estadístico I^2 como $I^2 = 100\% (Q - df)/Q$, en el que Q es el índice de heterogeneidad de Cochran y DF , grados de libertad. Un valor de 0% indica la inexistencia de heterogeneidad y los valores mayores indican la existencia de heterogeneidad. En aquellos análisis en los que el I^2 fue menor del 50%, se utilizó el modelo de efectos fijos y, en los que el I^2 fue mayor del 50%, se utilizó el modelo de efectos aleatorios. La validez de criterio general fue establecida mediante la correlación ponderada entre el EP y la FC, así como también entre el EP y el VO_2 . El análisis estadístico fue realizado

con el paquete estadístico MedCalc Statistical Software v. 14.12.0 (MedCalc Software bvba, Ostend, Belgium) y un intervalo de confianza de 95% (IC 95%) fue considerado estadísticamente significativo.

RESULTADOS

Selección de los estudios: La búsqueda arrojó 3338 artículos desde las 7 bases de datos. En la Figura 1, se describe la secuencia de búsqueda sistemática y las razones de la exclusión. Un total de 25 artículos se consideraron para la revisión (Figura 1).

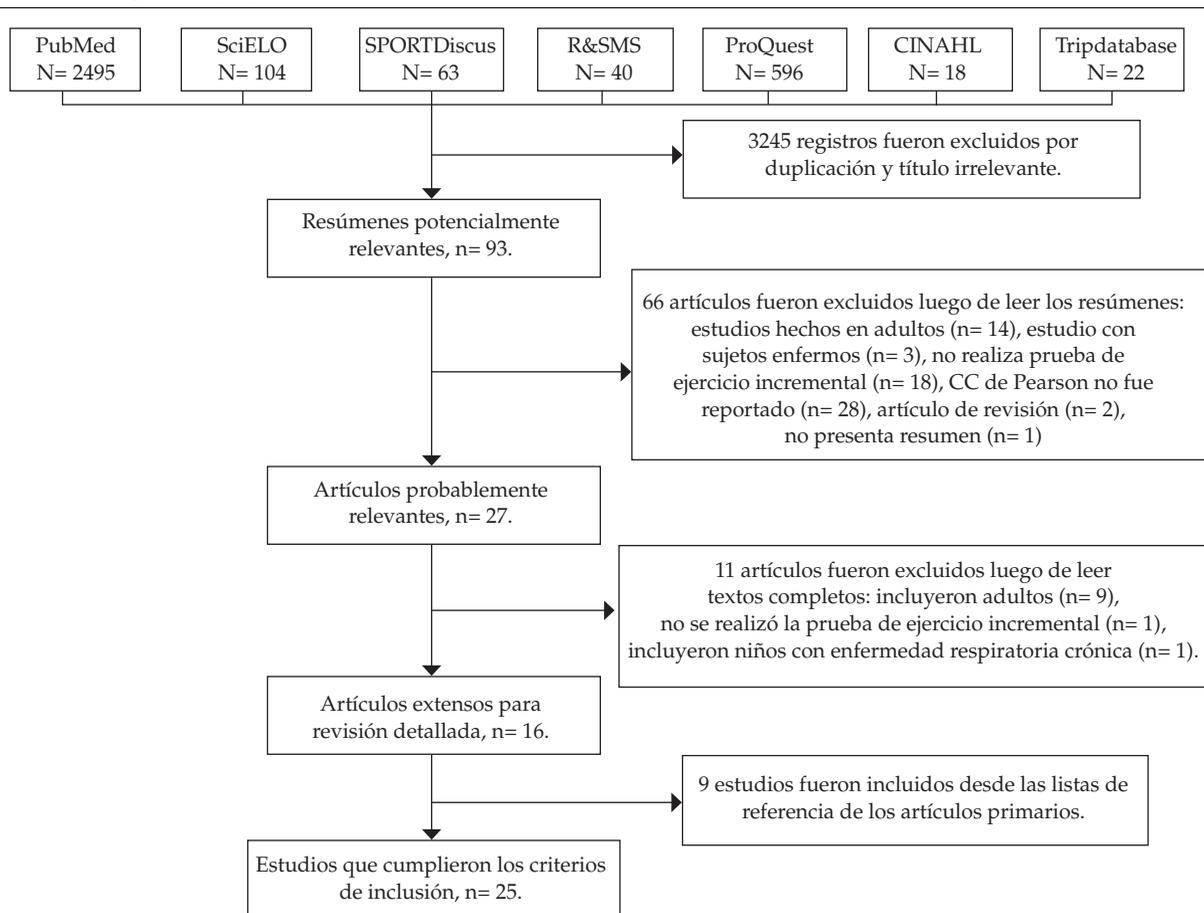
Características de los estudios: Todos los artículos fueron escritos en inglés. El rango de años de publicación fue 1986 y 2014. El tamaño muestral encontrado en los estudios fue entre 283 y 14 participantes; 1418 sujetos completaron el TEI en los 25 estudios, 699 (49,2%) fueron hombres y 721 (50,8%), mujeres. En 24 artículos (n= 1355;

95,5%), el promedio de edad fue reportado y, en 15 (n= 923; 65,0%), describieron los rangos de edad. El promedio ponderado de la edad fue 10,5 años (el promedio ponderado de la edad máxima y mínima fue 13,1 y 9,7 años). Por su parte, en 21 artículos (n= 1271; 89,6%), el protocolo de ejercicio fue continuo progresivo y, en 4 artículos (n= 147; 10,3%), fue intermitente progresivo. En 13 artículos, el modo de trabajo fue bicicleta (n= 831; 58,6%); en 9, carrera (n= 387; 27,2%); y 3 fueron prueba de escalón (n= 200; 14,1%).

Además, en 23 artículos, fue evaluada la validez bajo el paradigma de estimación (n= 1332; 93,9%) y 1 bajo el paradigma de producción (n= 70; 4,9%). Un artículo evaluó ambos paradigmas (n= 16; 1,1%).

En relación con el criterio de referencia utilizado, en 25 estudios, fue considerada la FC (n= 1418; 100%); en 16, VO₂ (n= 732; 51,6%); en 7, VE (n= 297; 20,9%); en 7, carga de trabajo (n= 302;

FIGURA 1. Diagrama de flujo



R&SMS: Fuentes de rehabilitación y medicina deportiva.
 CC: coeficiente de correlación.

21,2%); en 4, FR (n= 218; 15,3%); en 3, VCO_2/VO_2 (n= 155; 10,9%); y VE/VO_2 en 2 (n= 120; 8,4%).

Finalmente, en 4 estudios (5, 11-13), fueron reportados resultados de niños con alguna enfermedad concomitante y adultos. No obstante, solo los datos desde la población de interés fueron considerados (Tabla 1).

Calidad metodológica y riesgo de sesgo: Entre los evaluadores, fue calculado un CC intraclase (CCI) > 0,85. En los artículos discrepantes, se tomó una decisión basada en consenso. Así, la CM, en 6 estudios (24%), fue categorizada como “buena”; en 9 (36%), “razonable”; y, en 10 (40%), como “pobre”. En ningún estudio, la CM fue categorizada como “excelente”.

En relación con el riesgo de sesgo, se observó una distribución simétrica de los CC en el gráfico de embudo, y no existió concentración evidente de los estudios en ningún lado de la media ponderada (Anexo). Adicionalmente, se observó una débil correlación entre los CC y el tamaño muestral ($\tau = -0,148$; IC 95%: de -0,438 a 0,203). Esto indica la existencia de un bajo nivel de riesgo de sesgo en los estudios.

Validez de criterio de las escalas identificadas: En la búsqueda, fueron identificados 12 instrumentos de evaluación: (1) escala de categoría de 15 puntos de EP (*Rate of Perceived Exertion*; RPE 6-20, por sus siglas en inglés) (4-7, 12, 14, 15); (2) tabla de puntuación de esfuerzo de niños (*Children's Effort Rating Table*; CERT, por sus siglas en inglés) (6, 7, 14, 16-18); escala de esfuerzo percibido para niños OMNI y sus tres variaciones: (3) OMNI-bicicleta (11, 19-21), (4) OMNI-carrera (15, 22-25) y (5) OMNI-escalón (26); (6) escala de categoría de 15 puntos de EP para niños (*Rate of Perceived Exertion-Children*; RPE-C, por sus siglas en inglés);¹³ (7) tabla de puntuación de esfuerzo de niños con ilustraciones (*Pictorial Children's Effort Rating Table*; PCERT, por sus siglas en inglés);^{23,27,28} (8) escala de puntuación de esfuerzo de carro y cargas (*Cart and Load Effort Rating*; CALER, por sus siglas en inglés) (20); (9) escala de EP para niños (C-RPES);²⁹ (10) escala de puntuación de categorías de esfuerzo percibido de Borg (CR-10) (27); (11) escala curvilínea de puntuación de esfuerzo percibido de Eston-Parfitt (*Eston-Parfitt Ratings of Perceived Exertion*; EP-RPE, por sus siglas en inglés);^{30,31} (12) tarea de caída de canicas (MDT).^{30,31}

El modelo de efectos aleatorios mostró un CC ponderado entre el EP y las variables fisiológicas de 0,835 (IC 95%: 0,762-0,887) y 0,874 (IC 95%: 0,794-0,924) considerando la FC y el VO_2 como criterio de referencia (Figuras 2 y 3).

Además, solo fue posible desarrollar un metaanálisis específico en 7 de 12 escalas (RPE 6-20, CERT, OMNI-bicicleta, OMNI-marcha-carrera, PCERT, EP-RPE y MDT), en las cuales se observó que RPE 6-20 tuvo el menor nivel de correlación ponderado comparado a OMNI-bicicleta y EP-RPE (FC y VO_2). Adicionalmente, RPE 6-20 también mostró el menor CC ponderado comparado a CERT y PCERT considerando solo FC y VO_2 como criterio de referencia, respectivamente. No se observó una diferencia significativa en el resto de las escalas analizadas. Por su parte, en los estudios no metaanalizados, se consiguió verificar que RPE-C y CR-10 mostraron los menores CC ponderados y diferencia significativa con CALER y C-RPES. Solo CR-10 mostró menor nivel de validez que OMNI-escalón en el criterio VO_2 . En contraste, C-RPES mostró un mayor nivel de validez que CERT en el criterio VO_2 (Tabla 2).

En la Tabla 3, se ilustra la validez de criterio de acuerdo con las covariables. Solo fueron observadas diferencias significativas en el paradigma de evaluación y otros criterios de referencia, en los que tanto el criterio VE/VO_2 como el paradigma de producción mostraron un menor nivel de correlación estadísticamente significativo. El resto de las covariables mostraron un nivel de correlación de moderado a alto sin diferencia significativa entre sus categorías.

DISCUSIÓN

Este estudio mostró la existencia de una fuerte correlación global entre el EP y las variables fisiológicas (FC y VO_2) durante el ejercicio. Esto sugiere que la medición del EP sería válida en niños.

En relación con el análisis específico de las escalas identificadas, la RPE 6-20, RPE-C y CR-10 tuvieron los menores niveles de correlación. Vale la pena considerar que RPE 6-20 y CR-10 fueron desarrolladas para población adulta, por lo tanto, no poseen descriptores visuales intuitivos para niños.^{4,5,12,14,15,27} Muchos estudios han comparado la escala de Borg con escalas infantiles, y se ha mostrado un mejor rendimiento en las escalas adaptadas.^{6,14,15,27} Por su parte, RPE-C corresponde a una versión adaptada de la escala de Borg RPE 6-20 con ilustraciones visuales; sin embargo, solo ha mostrado correlación moderada con las variables fisiológicas durante el ejercicio.¹³ Aunque no ha sido comparada con otros instrumentos, es probable que RPE-C no

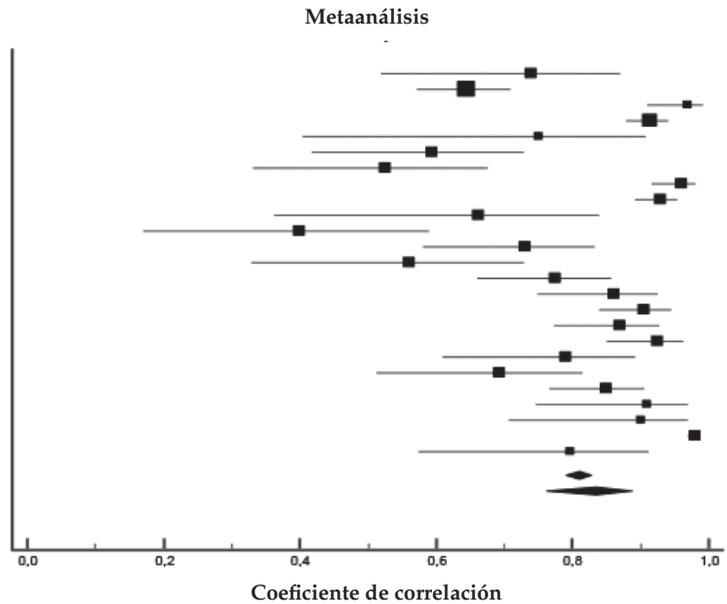
Tabla 1. Características de los estudios

Autor	Tamaño de muestra (M/F)	Edad	Protocolo de ejercicio	Modalidad de ejercicio	Paradigma de evaluación	Escala de EP	Criterio fisiológico	Calidad metodológica
Eston et al., 1986	30 (30/0)	16,0	Continuo progresivo	Bicicleta	Estimación	RPE 6-20	FC, carga	Razonable
Gillach et al., 1989	283 (144/139)	11,0	Continuo progresivo	Bicicleta	Estimación	RPE 6-20	FC	Razonable
Eakin et al., 1992	15 (7/8)	13,3	Continuo progresivo	Carrera	Estimación	RPE 6-20	FC, VO ₂	Pobre
William et al., 1994	112 (56/56)	7,1	Continuo progresivo	Escalón	Estimación	CERT	FC	Razonable
Eston et al., 1994	16 (8/8)	10,0	Continuo progresivo	Bicicleta	Estimación y producción	CERT	FC, carga	Pobre
Lamb, 1995	70 (28/42)	9,5	Continuo progresivo	Bicicleta	Estimación	RPE 6-20, CERT	FC, carga	Razonable
Lamb, 1996	70 (28/42)	9,5	Intermitente progresivo	Bicicleta	Producción	RPE 6-20, CERT	FC, carga	Razonable
Cassady et al., 1998	30 (17/13)	9,6	Continuo progresivo	Bicicleta	Estimación	C-RPES	FC, VO ₂ , FR, VE, VCO ₂ /VO ₂	Razonable
Robertson et al., 2000	80 (40/40)	10,0	Continuo progresivo	Bicicleta	Estimación	OMNI-bicicleta	FC, VO ₂	Buena
Gros Lambert et al., 2001	25 (13/12)	9,8	Continuo progresivo	Carrera	Estimación	RPE-C	FC	Razonable
Utter et al., 2002	63 (32/31)	13-6*	Continuo progresivo	Carrera	Estimación	OMNI-marcha-carrera	FC, VO ₂ , FR, VE, VE/VO ₂	Razonable
Pfeiffer et al., 2002	57 (0/57)	15,3	Continuo progresivo	Carrera	Estimación	RPE 6-20, OMNI-marcha-carrera	FC, VO ₂ , FR, VE, VE/VO ₂ , VCO ₂ /VO ₂	Buena
Yelling et al., 2002	48 (24/24)	13,8	Intermitente progresivo	Escalón	Estimación	PCERT	FC	Razonable
Leung et al., 2002	69 (34/35)	10,3	Continuo progresivo	Bicicleta	Estimación	RPE 6-20, CERT	FC, VO ₂ , carga	Buena
Robertson et al., 2005	40 (20/20)	11,1	Continuo progresivo	Escalón	Estimación	OMNI-escalón	FC, VO ₂	Razonable
Rommeich et al., 2006	51 (26/25)	11,2	Continuo progresivo	Carrera	Estimación	OMNI-marcha-carrera PCERT	FC, VO ₂	Buena
Robertson et al., 2006	44 (22/22)	12,8	Continuo progresivo	Carrera	Estimación	OMNI-marcha-carrera	FC, VO ₂	Razonable
Barkley et al., 2008	32 (16/16)	9,5	Continuo progresivo	Bicicleta	Estimación	OMNI-bicicleta, CALER	FC, VO ₂	Razonable
Marinov et al., 2008	50 (25/25)	10,4	Continuo progresivo	Carrera	Estimación	CR-10, PCERT	FC, VO ₂ , VE	Razonable
Suminski et al., 2008	68 (32/36)	10,2	Continuo progresivo	Carrera	Estimación	OMNI-marcha-carrera	FC, VO ₂ , FR, VCO ₂ /VO ₂ , VE	Buena
Leung et al., 2008	32 (17/15)	10,5	Continuo progresivo	Bicicleta	Estimación	CERT	FC, carga	Razonable
Eston et al., 2009	15 (6/9)	7,6	Intermitente progresivo	Bicicleta	Estimación	EP RPE, MDT	FC, VO ₂ , carga, VE	Pobre
Lambrick et al., 2011	14 (8/6)	7,9	Intermitente progresivo	Carrera	Estimación	EP RPE, MDT	FC, VO ₂ , VE	Pobre
Balasekaran et al., 2012	81 (45/36)	13,8	Continuo progresivo	Bicicleta	Estimación	OMNI-bicicleta	FC, VO ₂	Buena
Balasekaran et al., 2014	23 (23/0)	13,8	Continuo progresivo	Bicicleta	Estimación	OMNI-bicicleta	FC, VO ₂	Pobre

* Solo el rango es reportado; EP: esfuerzo percibido; escala de categoría de 15 puntos de EP (RPE 6-20); tabla de puntuación de esfuerzo de niños (CERT); escala de esfuerzo percibido para niños OMNI; escala de categoría de 15 puntos de EP para niños (RPE-C); tabla de puntuación de esfuerzo de niños con ilustraciones (PCERT); escala de puntuación de esfuerzo de carro y cargas (CALER); escala de EP para niños (C-RPES); escala de puntuación de categorías de esfuerzo percibido de Borg (CR-10); escala curvilínea de puntuación de esfuerzo percibido de Eston-Parfitt (EP-RPE); tarea de caída de canicas (MDT); FC: frecuencia cardíaca; VO₂: consumo de oxígeno; FR: frecuencia respiratoria; VE: ventilación minuto; VCO₂/VO₂: cociente respiratorio; VE/VO₂: equivalente ventilatorio para el oxígeno.

FIGURA 2. Metaanálisis de los estudios que consideran la frecuencia cardíaca como criterio de referencia

Estudio	n	r	IC 95%
Eston et al., 1986	30	0,740	0,518-0,869
Gillach et al., 1989	283	0,645	0,571-0,708
Eakin et al., 1992	15	0,970	0,910-0,990
William et al., 1994	112	0,915	0,879-0,941
Eston et al., 1994	16	0,750	0,405-0,908
Lamb, 1995	70	0,594	0,417-0,727
Lamb, 1996	70	0,525	0,331-0,677
Cassady et al., 1998	30	0,960	0,917-0,981
Robertson et al., 2000	80	0,930	0,893-0,955
Gros Lambert et al., 2001	25	0,663	0,363-0,838
Utter et al., 2002	63	0,400	0,169-0,589
Pfeiffer et al., 2002	57	0,730	0,580-0,832
Yelling et al., 2002	48	0,560	0,328-0,728
Leung et al., 2002	69	0,776	0,661-0,855
Robertson et al., 2005	40	0,860	0,749-0,924
Romneich et al., 2006	51	0,905	0,839-0,945
Robertson et al., 2006	44	0,870	0,773-0,927
Barkley et al., 2008	32	0,925	0,851-0,963
Leung et al., 2008	32	0,790	0,609-0,893
Marinov et al., 2008	50	0,692	0,513-0,814
Suminski et al., 2008	68	0,850	0,767-0,905
Eston et al., 2009	15	0,910	0,745-0,970
Lambrick et al., 2011	14	0,900	0,707-0,968
Balasekaran et al., 2012	81	0,980	0,969-0,987
Balasekaran et al., 2014	23	0,798	0,575-0,911
Total (efectos fijos)	1386	0,810	0,791-0,828
Total (efectos aleatorios)	1386	0,836	0,761-0,889

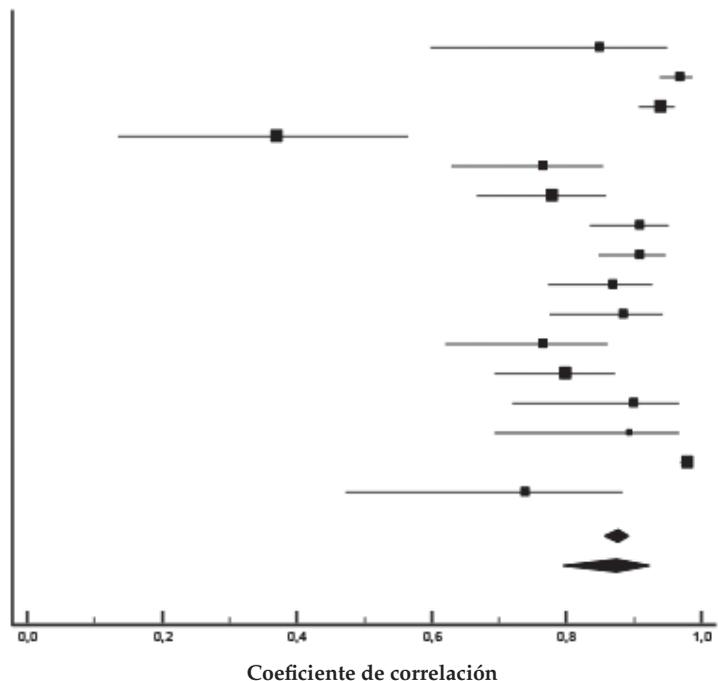


Prueba de heterogeneidad
 Q= 318,8134.
 DF= 24.
 Nivel de significancia= $p < 0,0001$.
 I² (inconsistencia)= 92,47%.
 IC95% I²= 90,08 a 94,29

IC 95%: intervalo de confianza; r: coeficiente de correlación ponderado; *se consideró el modelo de efectos aleatorios.
 Q: índice de heterogeneidad de Cochran; DF: grados de libertad.

FIGURA 3. Metaanálisis de los estudios que consideran el consumo de oxígeno como criterio de referencia

Estudio	n	r	IC 95%
Eakin et al., 1992	15	0,850	0,598-0,949
Cassady et al., 1998	30	0,970	0,937-0,986
Robertson et al., 2000	80	0,940	0,908-0,961
Utter et al., 2002	63	0,370	0,135-0,566
Pfeiffer et al., 2002	57	0,765	0,630-0,855
Leung et al., 2002	69	0,780	0,666-0,858
Robertson et al., 2005	40	0,910	0,835-0,952
Romneich et al., 2006	51	0,910	0,847-0,948
Robertson et al., 2006	44	0,870	0,773-0,927
Barkley et al., 2008	32	0,885	0,776-0,943
Marinov et al., 2008	50	0,766	0,621-0,861
Suminski et al., 2008	68	0,800	0,694-0,872
Eston et al., 2009	15	0,900	0,719-0,967
Lambrick et al., 2011	14	0,895	0,694-0,967
Balasekaran et al., 2012	81	0,980	0,969-0,987
Balasekaran et al., 2014	23	0,740	0,858-0,892
Total (efectos fijos)	732	0,876	0,857-0,892
Total (efectos aleatorios)	732	0,874	0,794-0,924



Q: índice de heterogeneidad de Cochran;
 DF: grados de libertad.

Prueba de heterogeneidad
 Q= 180,2103.
 DF= 15.
 Nivel de significancia= $p < 0,0001$.
 I² (inconsistencia)= 91,68%.
 IC95% I²= 88,09 a 94,18

IC 95%: intervalo de confianza; r: coeficiente de correlación ponderado;
 *se consideró el modelo de efectos aleatorios.

supere la validez de otras escalas para niños dado su complejo rango de categorías (de seis a veinte). Además, las escalas OMNI (bicicleta, marcha-carrera y escalón), EP-RPE, C-RPES, MDT, CALER y PCERT mostraron los CC más altos. Estas escalas incorporan números, descriptores verbales e ilustraciones adecuados para niños, lo cual puede explicar el alto nivel de validez mostrado.^{11,15,19-24,26,29-34}

En relación con el análisis de covariables, en nuestro estudio, se observó que el sexo, la edad, el protocolo y la modalidad de ejercicio no afectaron la validez de la medición del EP. Nuestros resultados discrepan de la RS realizada por Chen y cols., quienes observaron que algunas covariables podrían afectar la validez del EP.³ Sin embargo, en el estudio citado, fueron considerados adultos, sujetos con patología concomitante y otros tipos de ejercicio

no estandarizado, lo cual podría explicar estas diferencias.

Muchas líneas de evidencia muestran la existencia de una fuerte asociación entre el período del desarrollo y la habilidad para expresar el EP en niños.^{2,32,34} Recientemente, Rice y cols., observaron que la validez de la medición del EP se incrementaba proporcionalmente con la edad.³⁴ En nuestro estudio, los niños menores de 7 años mostraron menor correlación que los niños de entre 13 y 15 años, lo que sugiere la posible asociación entre madurez cognitiva y validez de la medición de EP. Además, el modo de producción mostró un menor nivel de validez que el modo de estimación ($p < 0,05$). Sin embargo, solo fueron identificados dos estudios primarios que evaluaron este paradigma mediante la correlación entre el EP y los parámetros fisiológicos.^{4,7} Se deja fuera de la RS aquellos estudios que emplearon

Tabla 2. Validez de criterio de las escalas identificadas

Escala	Criterio	Promedio r	k	n	IC 95%	p	Q	I ² (valor p)
RPE (6-20)	FC	0,697 ^a	7	523	0,649-0,739	< 0,001	11,3393	47,09% (0,0078)
	VO ₂	0,695 ^a	3	106	0,578-0,785	< 0,001	2,2748	12,08% (0,3207)
CERT	FC	0,778 ^b	6	301	0,572-0,892	< 0,001	49,1602	89,83% (< 0,0001)
	VO ₂	0,870	1	35	0,756-0,933	NA	NA	NA
OMNI-bicicleta	FC	0,935 ^b	4	216	0,843-0,974	< 0,001	30,9730	90,31% (< 0,0001)
	VO ₂	0,926 ^b	4	216	0,804-0,973	< 0,001	38,0031	91,11% (< 0,0001)
OMNI-marcha-carrera	FC	0,806 ^b	5	283	0,637-0,901	0,001	36,3875	89,01% (< 0,0001)
	VO ₂	0,819 ^b	5	283	0,627-0,917	< 0,001	48,0952	91,68% (< 0,0001)
OMNI-escalón	FC	0,860	1	40	0,749-0,924	NA	NA	NA
	VO ₂	0,910	1	40	0,835-0,952	NA	NA	NA
RPE-C	FC	0,663	1	25	0,363-0,838	NA	NA	NA
	VO ₂	-	-	-	-	-	-	-
PCERT	FC	0,769 ^b	3	149	0,482-0,907	< 0,001	17,6863	88,69% (0,0001)
	VO ₂	0,861 ^a	2	101	0,798-0,904	< 0,001	1,5549	35,69% (0,2124)
CALER	FC	0,920	1	32	0,841-0,961	NA	NA	NA
	VO ₂	0,880	1	32	0,767-0,940	NA	NA	NA
C-RPES	FC	0,960	1	30	0,917-0,981	NA	NA	NA
	VO ₂	0,970	1	30	0,937-0,986	NA	NA	NA
CR-10	FC	0,634	1	50	0,432-0,775	NA	NA	NA
	VO ₂	0,710	1	50	0,538-0,825	NA	NA	NA
EP-RPE	FC	0,916 ^a	2	29	0,820-0,962	< 0,001	0,7532	0,0% (0,3854)
	VO ₂	0,906 ^a	2	29	0,798-0,957	< 0,001	0,6074	0,0% (0,4358)
MDT	FC	0,866 ^a	2	29	0,721-0,939	< 0,001	0,08211	0,0% (0,7745)
	VO ₂	0,870 ^a	2	29	0,728-0,940	< 0,001	0,0000	0,0% (1,0)

k: número de artículos; n: tamaño de muestra; Q: índice de heterogeneidad; I²: índice de inconsistencia; NA: no fue realizado metaanálisis; ^a: se utilizó el modelo de efectos fijos; ^b: se utilizó el modelo de efectos aleatorios; escala de categoría de 15 puntos de EP (RPE 6-20); tabla de puntuación de esfuerzo de niños (CERT); escala de esfuerzo percibido para niños OMNI; escala de categoría de 15 puntos de EP para niños (RPE-C); tabla de puntuación de esfuerzo de niños con ilustraciones (PCERT); escala de puntuación de esfuerzo de carro y cargas (CALER); escala de EP para niños (C-RPES); escala de puntuación de categorías de esfuerzo percibido de Borg (CR-10); escala curvilínea de puntuación de esfuerzo percibido de Eston-Parfitt (EP-RPE); tarea de caída de canicas (MDT); FC: frecuencia cardíaca; VO₂: consumo de oxígeno; VE: ventilación minuto; VCO₂/VO₂: cociente respiratorio; VE/VO₂: equivalente ventilatorio para el oxígeno.

otros abordajes estadísticos, como el análisis de varianza (*analysis of variance*; ANOVA, por sus siglas en inglés), los cuales han mostrado resultados discrepantes.^{33,35} probablemente debido al complejo proceso psicofísico necesario para producir intensidad de ejercicio a partir del EP.² Futuros estudios deben ser realizados para verificar todas estas hipótesis.

En relación con el análisis de los criterios de referencia, se observó que, excepto el VE/VO₂ (Tabla 3), las otras variables también serían criterios útiles para la determinación de la validez de criterio del EP, al igual que la FC y el VO₂.

Entre las limitaciones de esta RS, se debe mencionar que 18 estudios (75%) tuvieron una calidad metodológica “razonable” o “pobre”. Las principales razones son asociadas con debilidades en el diseño metodológico y escaso tamaño muestral, lo cual afecta la potencia estadística de los resultados. Además, una alta heterogeneidad fue observada, probablemente

debido a las diferencias metodológicas entre los estudios. En este contexto, las diferencias en el tamaño muestral, diseño del protocolo incremental, en términos de la duración y cantidad de los intervalos de intensidad, así como también las diferencias en la metodología empleada para la medición del EP, fueron observadas en los estudios primarios. Esto pudo haber contribuido a los altos niveles de heterogeneidad observados.

Dado el alto nivel de validez externa de este diseño de investigación, los resultados obtenidos permiten respaldar la utilización del EP para monitorizar la intensidad desarrollada durante el ejercicio en sujetos menores de 18 años. Pese a lo anterior, aún es necesario verificar sus propiedades psicométricas en niños con patología concomitante y en otras modalidades de ejercicio.

Por último, los resultados de este estudio sugieren que la medición del EP sería válida en niños sanos durante una prueba de ejercicio

TABLA 3. Validez de criterio de la medición del esfuerzo percibido

Covariables	Promedio r	k	n	IC 95%	p	Q	I ² (valor p)
Sexo							
Varón	0,847 ^b	14	370	0,780-0,895	< 0,001	75,8810	73,64% (< 0,0001)
Niña	0,851 ^b	13	391	0,774-0,903	< 0,001	117,9435	83,04% (< 0,0001)
* Edad							
< 7 años	0,821 ^b	2	41	0,450-0,950	0,001	7,5838	73,63% (0,0226)
8-12 años	0,863 ^b	18	1040	0,803-0,906	< 0,001	285,2875	90,89% (< 0,0001)
13-15 años	0,937 ^b	4	131	0,828-0,978	< 0,001	39,9800	87,49% (< 0,0001)
> 15 años	0,723 ^b	3	81	0,493-0,858	< 0,001	8,6094	76,77% (0,0135)
Prueba de ejercicio							
Continuo progresivo	0,875 ^b	21	1271	0,812-0,918	< 0,001	414,6569	93,01% (< 0,0001)
Intermitente progresivo	0,760 ^b	4	147	0,602-0,860	< 0,001	37,7057	81,44% (< 0,0001)
Modalidad de ejercicio							
Bicicleta	0,865 ^b	13	831	0,784-0,917	< 0,001	273,6051	92,69% (< 0,0001)
Carrera	0,807 ^b	9	387	0,722-0,867	< 0,001	66,4815	80,45% (< 0,0001)
Escalón	0,907 ^b	3	200	0,704-0,973	< 0,001	140,7281	95,74% (< 0,0001)
Paradigma de evaluación							
Estimación	0,868 ^b	23	1348	0,815-0,906	< 0,001	452,3205	91,60% (< 0,0001)
Producción	0,550 ^a	2	86	0,427-0,653	< 0,001	2,3499	14,89% (0,3088)
Otros criterios de referencia							
Carga de trabajo	0,811 ^b	7	303	0,724-0,873	< 0,001	43,0231	75,01% (< 0,0001)
FR	0,695 ^b	4	218	0,437-0,847	< 0,001	40,4234	90,10% (< 0,0001)
VE	0,828 ^b	7	297	0,698-0,906	< 0,001	98,2091	89,82% (< 0,0001)
VCO ₂ /VO ₂	0,732 ^b	3	155	0,556-0,845	< 0,001	14,2162	78,90% (0,0026)
VE/VO ₂	0,260	2	120	-0,0103-0,495	0,059	6,6834	70,08% (0,0354)

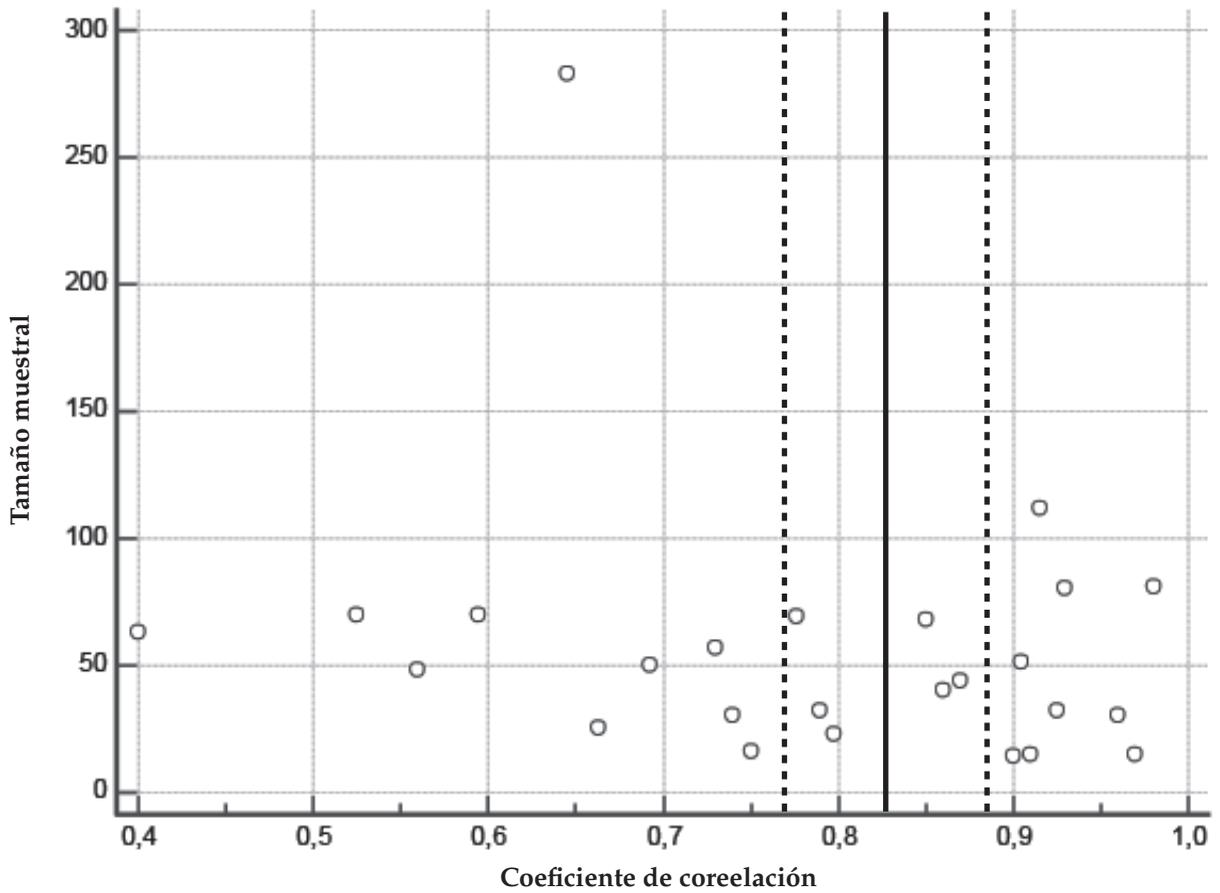
k: número de artículos; n: tamaño de muestra; Q: índice de heterogeneidad; I²: índice de inconsistencia; NA: no fue realizado metaanálisis; ^a: se utilizó el modelo de efectos fijos; ^b: se utilizó el modelo de efectos aleatorios; escala de categoría de 15 puntos de EP (RPE 6-20); tabla de puntuación de esfuerzo de niños (CERT); escala de esfuerzo percibido para niños OMNI; escala de categoría de 15 puntos de EP para niños (RPE-C); tabla de puntuación de esfuerzo de niños con ilustraciones (PCERT); escala de puntuación de esfuerzo de carro y cargas (CALER); escala de EP para niños (C-RPES); escala de puntuación de categorías de esfuerzo percibido de Borg (CR-10); escala curvilínea de puntuación de esfuerzo percibido de Eston-Parfitt (EP-RPE); tarea de caída de canicas (MDT); FC: frecuencia cardíaca; VO₂: consumo de oxígeno; FR: frecuencia respiratoria; VE: ventilación minuto; VCO₂/VO₂: cociente respiratorio; VE/VO₂: equivalente ventilatorio para el oxígeno.

incremental. Además, las escalas específicamente adaptadas para niños parecen tener mejor rendimiento que aquellas no adaptadas. Sin embargo, estudios con mejor calidad metodológica deben ser realizados para confirmar estas conclusiones. ■

REFERENCIAS

- Garber CE, Blissmer B, Deschenes MR, Franklin BA, 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 Sports Exerc* 2011;43(7):1334-59.
- Gros Lambert A, Mahon AD. Perceived exertion: influence of age and cognitive development. *Sports Med* 2006;36(11):911-28.
- Chen MJ, Fan X, Moe ST. Criterion-related validity of the Borg ratings of perceived exertion scale in healthy individuals: a meta-analysis. *J Sports Sci* 2002;20(11):873-99.
- Eston RG, Williams JG. Exercise intensity and perceived exertion in adolescent boys. *Br J Sports Med* 1986;20(1):27-30.
- Gillach MC, Sallis JF, Buono MJ, Patterson P, et al. The relationship between perceived exertion and heart rate in children and adults. *Pediatr Exerc Sci* 1989;1(4):360-8.
- Lamb KL. Children's ratings of effort during cycle ergometry: an examination of the validity of two effort rating scales. *Pediatr Exerc Sci* 1995;7(4):407-21.
- Lamb KL. Exercise regulation during cycle ergometry using the Children's Effort Rating Table (CERT) and Rating of Perceived Exertion (RPE) scales. *Pediatr Exerc Sci* 1996;8(4):337-50.
- Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg* 2010;8(5):336-41.
- Mokkink LB, Terwee CB, Patrick DL, Alonso J, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63(7):737-45.
- Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;343:d4002.
- Balasekaran G, Loh MK, Govindaswamy VV, Cai SJ. Omni Scale Perceived Exertion responses in obese and normal weight male adolescents during cycle exercise. *J Sports Med Phys Fitness* 2014;54(2):186-96.
- Eakin BL, Finta KM, Serwer GA, Beekman RH. Perceived exertion and exercise intensity in children with or without structural heart defects. *J Pediatr* 1992;120(1):90-3.
- Gros Lambert A, Hintzy F, Hoffman MD, Dugué B, et al. Validation of a rating scale of perceived exertion in young children. *Int J Sports Med* 2001;22(2):116-9.
- Leung ML, Chung PK, Leung RW. An assessment of the validity and reliability of two perceived exertion rating scales among Hong Kong children. *Percept Mot Skills* 2002;95(3 Pt 2):1047-62.
- Pfeiffer KA, Pivarnik JM, Womack CJ, Reeves MJ, et al. Reliability and validity of the Borg and OMNI rating of perceived exertion scales in adolescent girls. *Med Sci Sports Exerc* 2002;34(12):2057-61.
- Eston RG, Lamb KL, Bain A, Williams AM, et al. Validity of a perceived exertion scale for children: a pilot study. *Percept Mot Skills* 1994;78(2):691-7.
- Williams JG, Eston R, Furlong B. CERT: a perceived exertion scale for young children. *Percept Mot Skills* 1994;79(3 Pt 2):1451-8.
- Leung R, Pak-Kwong C, Bower G. Validation of a translated children perceived exertion rating scale. *Int J Fit* 2008;4(2):49-56.
- Balasekaran G, Loh MK, Govindaswamy VV, Robertson RJ. OMNI Scale of Perceived Exertion: mixed gender and race validation for Singapore children during cycle exercise. *Eur J Appl Physiol* 2012;112(10):3533-46.
- Barkley JE, Roemmich JN. Validity of the CALER and OMNI-bike ratings of perceived exertion. *Med Sci Sports Exerc* 2008;40(4):760-6.
- Robertson RJ, Goss FL, Boer NF, Peoples JA, et al. Children's OMNI scale of perceived exertion: mixed gender and race validation. *Med Sci Sports Exerc* 2000;32(2):452-8.
- Robertson RJ, Goss FL, Aaron DJ, Tessmer KA, et al. Observation of perceived exertion in children using the OMNI pictorial scale. *Med Sci Sports Exerc* 2006;38(1):158-66.
- Roemmich JN, Barkley JE, Epstein LH, Lobarinas CL, et al. Validity of PCERT and OMNI walk/run ratings of perceived exertion. *Med Sci Sports Exerc* 2006;38(5):1014-9.
- Suminski RR, Robertson RJ, Goss FL, Olvera N. Validation of the Omni Scale of Perceived Exertion in a sample of Spanish-speaking youth from the USA. *Percept Mot Skills* 2008;107(1):181-8.
- Utter AC, Robertson RJ, Nieman DC, Kang J. Children's OMNI Scale of Perceived Exertion: walking/running evaluation. *Med Sci Sports Exerc* 2002;34(1):139-44.
- Robertson RJ, Goss FL, Andreacci JL, Dubé JJ, et al. Validation of the children's OMNI RPE scale for stepping exercise. *Med Sci Sports Exerc* 2005;37(2):290-8.
- Marinov B, Mandadjieva S, Kostianev S. Pictorial and verbal category-ratio scales for effort estimation in children. *Child Care Health Dev* 2008;34(1):35-43.
- Yelling M, Lamb KL, Swaine IL. Validity of a pictorial perceived exertion scale for effort estimation and effort production during stepping exercise in adolescent children. *Eur Phy Educ Rev* 2002;8(2):157-75.
- Cassady SL, Kaufman BA, Kelly CE, Eisenmann SC, et al. Validity of a New Perceived Exertion Scale for Children. *Cardiopulm Phys Ther J* 1998;9(1):3-8.
- Eston RG, Lambrick DM, Rowlands AV. The perceptual response to exercise of progressively increasing intensity in children aged 7-8 years: validation of a pictorial curvilinear ratings of perceived exertion scale. *Psychophysiology* 2009;46(4):843-51.
- Lambrick DM, Rowlands AV, Eston RG. The perceptual response to treadmill exercise using the Eston-Parfitt scale and marble dropping task, in children age 7 to 8 years. *Pediatr Exerc Sci* 2011;23(1):36-48.
- Gros Lambert A, Monnier Benoit P, Grange CC, Rouillon JD. Self-regulated running using perceived exertion in children. *J Sports Med Phys Fitness* 2005;45(1):20-5.
- Parfitt G, Shepherd P, Eston RG. Reliability of effort production using the children's CALER and BABE perceived exertion scales. *J Exerc Sci Fit* 2007;5(1):49-55.
- Rice KR, Gammon C, Pfeiffer K, Trost SG. Age related differences in the validity of the OMNI perceived exertion scale during lifestyle activities. *Pediatr Exerc Sci* 2015;27(1):95-101.
- Ward DS, Jackman JD, Galiano FJ. Exercise intensity reproduction: children versus adults. *Pediatr Exerc Sci* 1991;3(3):209-18.

Anexo. Gráfico de embudo



La línea continua muestra la media ponderada; la línea segmentada, 95% de intervalo de confianza.