

Validez concurrente y confiabilidad interobservador de la escala de medición de esfuerzo percibido infantil EPInfant en niños sanos chilenos

Concurrent validity and interobserver reliability of the EPInfant pediatric perceived exertion rating scale among healthy Chilean children

Klgo. Iván Rodríguez-Núñez^{a,b,c} y Dr. Carlos Manterola.^{c,d,e}

RESUMEN

El objetivo de este estudio fue determinar la validez y la confiabilidad de la escala de medición de esfuerzo percibido infantil EPInfant en niños chilenos. Fueron seleccionados niños de entre 8 y 12 años, en quienes se registró frecuencia cardíaca, esfuerzo percibido y carga de trabajo durante dos tests de Chester consecutivos con una semana de intervalo. Se calcularon el coeficiente *r* de Pearson y el coeficiente de correlación intraclase. Al estudio, ingresaron 31 niños (17 varones). Existió una fuerte correlación entre el esfuerzo percibido y los criterios de referencia, así como también se calculó un alto coeficiente de correlación intraclase. El promedio de la diferencia entre el esfuerzo percibido en ambos tests fue 1,4 (95% IC: de -1,4 a 4,1) y 0,2 (95% IC: de -1,8 a 2,2) en varones y niñas, respectivamente. La escala EPInfant muestra una validez de criterio y confiabilidad interobservador aceptable en niños durante una prueba de ejercicio incremental en escalón.

Palabras clave: confiabilidad y validez, niños, escalas, ejercicio, esfuerzo físico.

INTRODUCCIÓN

Las escalas de medición del esfuerzo percibido (EP) permiten cuantificar la sensación causada

por los cambios metabólicos durante el ejercicio. Así, la escala de Borg ha sido ampliamente utilizada y ha mostrado niveles aceptables de validez en adulto.¹ Sin embargo, no existe evidencia favorable respecto a su rendimiento en población infantil.^{2,3}

Recientemente, se ha descrito que las propiedades psicométricas de estas escalas en niños dependen, en gran medida, del desarrollo cognitivo del sujeto y del tipo de instrumento utilizado.⁴ Así, fue desarrollada la escala de medición de esfuerzo percibido infantil (EPInfant). Esta posee descriptores numéricos (de 0 a 10), descriptores verbales en español y un set de ilustraciones que representan a un niño corriendo a lo largo de una escala creciente.⁵ Ha mostrado un alto nivel de validez en adolescentes de entre 13 y 15 años.^{5,6} No obstante, ha sido escasamente evaluada en niños menores.⁷

Por lo tanto, el objetivo de este estudio fue determinar la validez de criterio concurrente y la confiabilidad interobservador de la escala EPInfant en niños sanos de entre 8 y 12 años de edad durante una prueba de ejercicio incremental.

MÉTODOS

El presente estudio posee un diseño test-retest, en el cual fueron seleccionados niños sanos de entre 8 y 12 años de un colegio de Concepción, Chile. Se excluyeron del estudio niños obesos, con alteración cognitiva, enfermedad cardiorrespiratoria o neuromotora. Para el cálculo de tamaño muestral, se consideró un coeficiente de correlación intraclase (CCI) de 0,7, riesgo de error α 5%, error β 10% y un intervalo de confianza (IC) para el CCI inferior a 0,15. Así, la muestral mínima necesaria para este estudio fueron 14 varones y 14 niñas.

Protocolo de estudio: Una semana antes del inicio de las pruebas, se registró la edad, el género y las variables antropométricas. El peso y la estatura fueron determinadas mediante una balanza análoga y cinta métrica. En esa misma

-
- Escuela de Kinesiología, Facultad de Ciencias de la Salud, Universidad San Sebastián, Concepción, Chile.
 - Carrera de Kinesiología, Facultad de Salud, Universidad Santo Tomás, Concepción, Chile.
 - Doctorado en Ciencias Médicas, Facultad de Medicina, Universidad de la Frontera, Temuco, Chile.
 - Departamento de Cirugía, Universidad de la Frontera, Chile.
 - Centro de Investigación en Ciencias Biomédicas, Universidad Autónoma de Chile.

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

Financiamiento: este estudio fue financiado por la Escuela de Kinesiología, Facultad de Ciencias de la Salud, Universidad San Sebastián, Concepción, Chile.

Conflicto de intereses: ninguno que declarar.

Recibido: 27-9-2015
Aceptado: 27-1-2016

instancia, los participantes se familiarizaron con la escala EPInfant de acuerdo con lo publicado por Rodríguez y cols.⁵

El test de ejercicio correspondió al test de Chester (TChester).⁸ Este se realizó en dos ocasiones con una semana de intervalo. El TChester consiste en subir y bajar un escalón de 20 cm a una velocidad incremental proporcionada por una señal sonora. Posee 5 niveles de acuerdo con el número de ascensos y descensos con ambos pies que el sujeto puede desarrollar en un minuto. Cada nivel dura 2 minutos y su duración total es 10 minutos.

Antes del TChester, los niños se mantuvieron en reposo por 10 minutos, se registró la frecuencia cardíaca (FC) (Monitor Polar®) y se calculó la FC máxima mediante la fórmula de Tanaka: $208 - (0,7 \times \text{edad})$.

Durante el TChester, la FC y el EP fueron registrados por dos evaluadores de modo independiente durante los últimos 20 segundos de cada minuto. Adicionalmente, la medición del EP en la evaluación 1 y 2 se llevó a cabo por dos evaluadores distintos. Además, la carga de trabajo (*watts*) fue calculada de acuerdo con la siguiente fórmula: altura del escalón (m) \times número de ciclos \times peso (kg) \times 0,16357.⁹

Este estudio fue aprobado por el Comité de Ética de la Universidad San Sebastián y el cuidador legal de cada participante firmó un consentimiento informado. El estudio fue realizado entre los meses de septiembre y diciembre de 2014.

Análisis estadístico: Utilizando el *software* estadístico MedCalc 14.12, se realizó estadística descriptiva con cálculo de promedio y desviación estándar (DE), y se evaluó la normalidad mediante el test de Shapiro-Wilk. Fue utilizado el test de análisis de la varianza (*analysis of variance*; ANOVA, por sus siglas en inglés) de medidas repetidas con ajuste de Bonferroni para evaluar las diferencias en la FC, las cargas de trabajo y el EP entre los niveles de intensidad durante el TChester. Se utilizó la prueba de T-Student para muestras independientes para evaluar las diferencias entre sexos.

El análisis de validez de criterio concurrente se realizó mediante la regresión lineal entre el EP y la FC, así como también entre el EP y la carga de trabajo. Por su parte, para la confiabilidad interobservador, se calculó el CCI. Adicionalmente, fue calculado el promedio de las diferencias del EP y su límite de concordancia mediante el método de Bland-Altman. Los

resultados de validez y confiabilidad fueron expresados en valor absoluto y 95% de IC. Se consideró significativo un valor de $p < 0,05$.

RESULTADOS

Al estudio, ingresaron 31 niños (17 varones) de entre 8 y 12 años. Las características de la muestra son reportadas en la *Tabla 1*.

La intensidad de ejercicio durante el TChester fue significativamente mayor en varones que en niñas. Por su parte, la FC y el EP incrementaron de manera significativa en cada nivel de intensidad ($p < 0,05$). No se observó diferencia en los niveles máximos de FC y EP alcanzados entre varones y mujeres.

Solo en las mujeres, se advirtió una menor FC en el primer nivel del TChester en la primera evaluación y no se observaron diferencias en los otros niveles. Además, solo los varones presentaron, en todas las intensidades del TChester, un menor nivel de EP en la primera prueba.

En relación con la validez de criterio, existió una fuerte correlación positiva entre el EP y la FC, así como también entre el EP y la carga de trabajo (*Figura 1. A y B*). Solo hubo diferencia significativa entre géneros en la pendiente de la regresión entre el EP y la carga de trabajo ($p = 0,03328$). Finalmente, todos los modelos de regresión fueron estadísticamente significativos ($p < 0,0001$). En la *Tabla 2*, se muestran los componentes de las ecuaciones de regresión.

En relación con la confiabilidad, el CCI general fue de 0,88 (95% IC: 0,82-0,92) y 0,94 (95% IC: 0,91-0,96) para varones y niñas, respectivamente. En cada nivel del TChester, se observaron niveles de confiabilidad de moderados a fuertes, que alcanzaron significancia estadística en ambos

TABLA 1. Características generales de la muestra de estudio

	Varones (n= 17)	Niñas (n= 14)
Edad (años)	10,1 \pm 0,6	8,7 \pm 0,9
Peso (kg)	42,1 \pm 9,4	31,4 \pm 5,9
Estatura (cm)	144,9 \pm 11,0	133,5 \pm 7,4
FC máxima teórica (lat./min)	201,0 \pm 0,8	201,5 \pm 0,8
FC al final del TChester 1	160,6 \pm 15,6	155,5 \pm 14,6
FC al final del TChester 2	164,2 \pm 13,8	162,1 \pm 17,9
Carga de trabajo al final del TChester (<i>watts</i>)	47,9 \pm 10,3	35,5 \pm 6,1

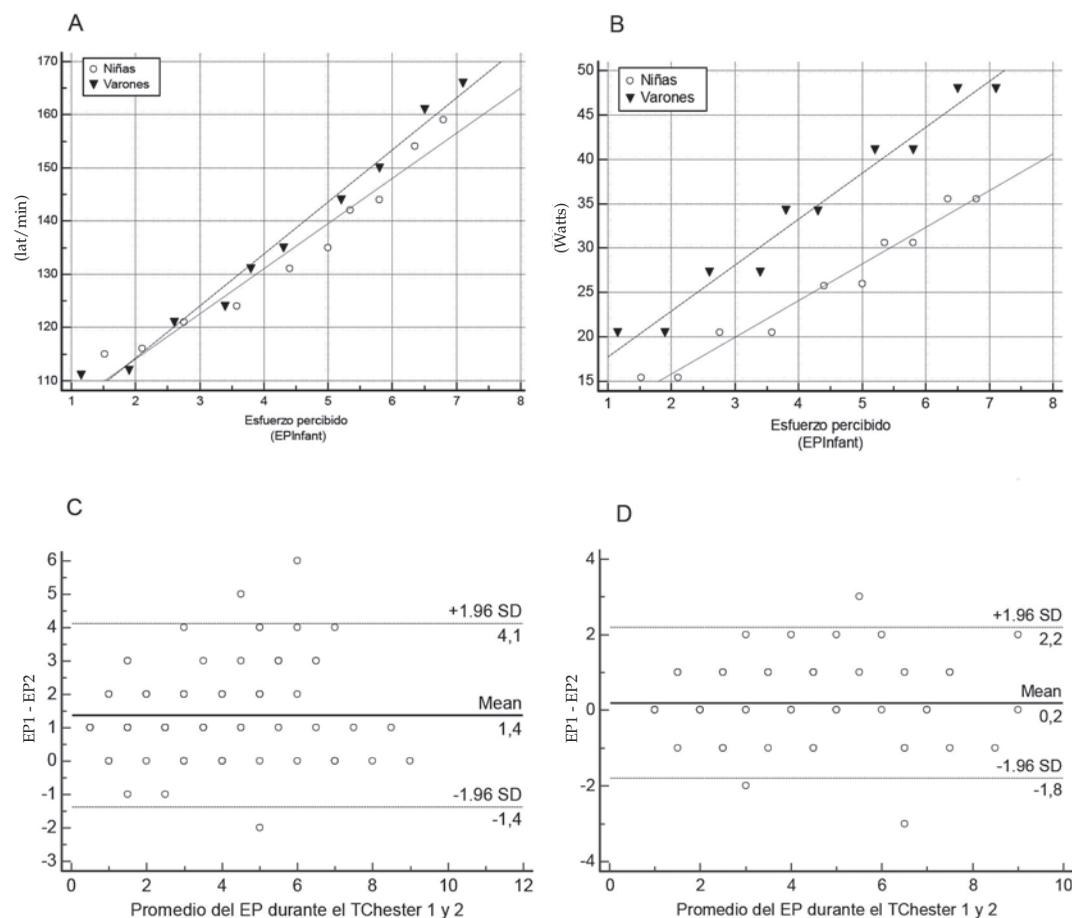
FC: frecuencia cardíaca; TChester: test de escalón incremental de Chester. Los resultados son expresados en promedio y desviación estándar.

géneros, excepto en el nivel I (Tabla 3). Además, la diferencia media entre las mediciones del EP fue 1,4 (95% IC: 4,1-1,4) y 0,2 (95% IC: 2,2-1,8) en varones y niñas, respectivamente (Figura 1. C y D).

DISCUSIÓN

Los principales hallazgos de este estudio confirman que la escala EPInfant posee una validez y una confiabilidad aceptables en niños de entre 8 y 12 años durante una prueba de ejercicio en escalón.

Figura 1. Validez de criterio y confiabilidad de la escala EPInfant. En A, se observa la correlación entre la frecuencia cardíaca y el esfuerzo percibido y, en B, la correlación entre la carga de trabajo y el esfuerzo percibido durante el TChester. C: gráfico de Bland-Altman de diferencias individuales en varones. D: gráfico de Bland-Altman de diferencias individuales en niñas.



EP: esfuerzo percibido.

Los resultados se muestran en promedio y 95% de intervalo de confianza.

Tabla 2. Análisis de regresión lineal de la escala de medición de esfuerzo percibido EPInfant expresado como función de la frecuencia cardíaca (lat./min) y el trabajo desarrollado durante el TChester (watts)

Variable	Subgrupo	Pendiente	Intercepto	r	r ²	Valor p
Frecuencia cardíaca	Todos	9,0	96,26	0,97	0,95	< 0,0001
	Varones	9,7	94,9	0,99	0,98	< 0,0001
	Niñas	8,3	97,7	0,97	0,95	< 0,0001
Carga de trabajo	Todos	4,4	10,7	0,84	0,71	< 0,0001
	Varones	5,0	12,9	0,98	0,96	< 0,0001
	Niñas	4,0	7,7	0,98	0,96	< 0,0001

Estos resultados concuerdan con los estudios de validez de otras escalas estudiadas en población anglosajona y en otras poblaciones no hispanoparlantes, las cuales han mostrado un alto nivel de validez durante el ejercicio incremental en niños (r de Pearson $> 0,85$).^{10,11} En nuestro estudio, los coeficientes de determinación (r^2) fueron superiores a 0,9 en ambos géneros, lo que demuestra el alto nivel de interdependencia entre los criterios de referencia y el EP (EPInfant) durante el ejercicio.

Además, la confiabilidad de estos instrumentos ha sido escasamente verificada en niños. Leung y cols. observaron un CCI superior a 0,85 en niños de entre 10 y 11 años durante una prueba de ejercicio incremental.² Pfeiffer y cols. advirtieron un aceptable nivel de confiabilidad en la medición del EP durante el ejercicio, principalmente en escalas adaptadas para población infantil (OMNI).³

Se debe destacar que el CCI representa la proporción de la variabilidad que no es causada por error de medición.¹² Así, considerando el alto CCI observado en ambos géneros, se confirma la existencia de un bajo riesgo de error aleatorio de medición en la aplicación de la escala EPInfant. Por su parte, en el análisis de confiabilidad para cada nivel del TChester, se observó un mayor CCI en los tres últimos niveles de intensidad de ejercicio, por lo tanto, es probable que la experimentación del estrés fisiológico haya disminuido la variabilidad en el nivel de EP reportado. Esto concuerda con algunos reportes que sugieren el rol sustancial de la experiencia en la capacidad perceptiva de esfuerzo de las personas.¹³

Entre las debilidades de este estudio, se debe considerar la diferencia de edades entre los géneros; no obstante, ambos grupos fueron

categorizados en la etapa de operaciones concretas según Piaget. Además, el escaso número de sujetos estudiado no permite extrapolar estos resultados, por lo cual este estudio debe ser reproducido en otras poblaciones.

En conclusión, la escala EPInfant mostró una adecuada validez de criterio y confiabilidad interobservador para la medición del EP en esta muestra de niños sanos de entre 8 y 12 años. Futuros estudios son necesarios para confirmar estos resultados. ■

Agradecimientos

A todos los niños y sus padres que participaron en este estudio.

REFERENCIAS

1. 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.
2. 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.
3. 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.
4. Rodríguez I, Zambrano L, Manterola C. Validez de criterio de las escalas de medición de esfuerzo percibido en niños sanos: una revisión sistemática y metaanálisis. *Arch Argent Pediatr* 2016;114(2):120-8.
5. Rodríguez I, Zenteno D, Cisternas L, Rodríguez P, et al. Construcción y evaluación de EPInfant: una escala para la medición del esfuerzo percibido en población pediátrica. *Arch Argent Pediatr* 2015;113(6):550-7.
6. Rodríguez I, Manterola C. Validación inicial de la escala de medición de esfuerzo percibido infantil (EPInfant) en niños chilenos. *Biomédica* 2016;36(1). [Epub 2015 21 Jul].
7. Rodríguez Nuñez I. Escala de medición de esfuerzo percibido infantil (EPInfant): validación en niños y adolescentes chilenos. *Rev Chil Pediatr* 2015. [Epub 2015 Oct 8].

TABLA 3. Confiabilidad de la escala de medición de esfuerzo percibido EPInfant durante el TChester

Varones	Nivel I	Nivel II	Nivel III	Nivel IV	Nivel V
EP 1	2,0 ± 1,1	3,4 ± 1,5	4,7 ± 1,9	5,5 ± 1,6	6,5 ± 1,8
EP 2	1,1 ± 0,8*	2,1 ± 1,3*	2,8 ± 1,6*	4,1 ± 1,7*	5,2 ± 1,7*
CCI	0,58	0,67*	0,77*	0,84*	0,73*
Niñas	Nivel I	Nivel II	Nivel III	Nivel IV	Nivel V
EP 1	1,5 ± 0,6	2,6 ± 1,2	3,8 ± 1,5	5,2 ± 1,6	6,6 ± 2,1
EP 2	1,4 ± 0,6	2,6 ± 1,2	3,6 ± 1,6	4,8 ± 1,6	6,4 ± 1,9
CCI	0,69	0,76*	0,92*	0,89*	0,84*

EP: esfuerzo percibido; TChester: test de escalón incremental de Chester; CCI: coeficiente de correlación intraclase;

* $p < 0,05$; los resultados se muestran en promedio y desviación estándar.

* Los varones presentaron diferencia significativa en el EP entre la evaluación 1 y la evaluación 2 (nivel 1, $p = 0,002$; nivel 2, $p = 0,0016$; nivel 3, $p = 0,0002$; nivel 4, $p = 0,0003$; nivel 5, $p = 0,004$. Se utilizó T-student para muestras pareadas).

8. Buckley JP, Sim J, Eston RG, Hession R, et al. Reliability and validity of measures taken during the Chester step test to predict aerobic power and to prescribe aerobic exercise. *Br J Sports Med* 2004;38(2):197-205.
9. De Andrade CH, de Camargo AA, de Castro BP, Malaguti C, et al. Comparison of cardiopulmonary responses during 2 incremental step tests in subjects with COPD. *Respir Care* 2012;57(11):1920-6.
10. Barkley JE, Roemmich JN. Validity of the CALER and OMNI-bike ratings of perceived exertion. *Med Sci Sports Exerc* 2008;40(4):760-6.
11. Gros Lambert A, Mahon AD. Perceived exertion: influence of age and cognitive development. *Sports Med* 2006;36(11):911-28.
12. Zou GY. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. *Stat Med* 2012;31(29):3972-81.
13. Barroso R, Cardoso RK, do Carmo EC, Tricoli V. Perceived exertion in coaches and young swimmers with different training experience. *Int J Sports Physiol Perform* 2014;9(2):212-6.