



Artículo original

# Confiabilidade do teste de corrida/caminhada de 9 minutos em crianças e adolescentes de 7-12 anos de idade



A.C. Paludo<sup>a,c,d,\*</sup>, M.B. Batista<sup>b,c,d</sup>, H. Serassuelo Junior<sup>b,c,d</sup>, G.B. Shigaki<sup>b,c,d</sup>, E.S. Cyrino<sup>b,c,d</sup> e E.R.V. Ronque<sup>b,c,d</sup>

<sup>a</sup> Universidade de São Paulo USP, Programa de Pós Graduação em Educação Física, São Paulo, Brasil

<sup>b</sup> Universidade Estadual de Londrina UEL, Programa de Pós-Graduação Associado em Educação Física UEL-UEM, Londrina, Brasil

<sup>c</sup> Grupo de Estudo e Pesquisa em Atividade Física e Exercício- GEPAFE.CEFE, Universidade Estadual de Londrina (UEL), Londrina, PR, Brasil

<sup>d</sup> Grupo de Estudo e Pesquisa em Metabolismo, Nutrição e Exercício GEPEMENE. CEFE, Universidade Estadual de Londrina (UEL), Londrina, PR, Brasil

## INFORMAÇÃO SOBRE O ARTIGO

### Historial do artigo:

Recebido a 2 de junho de 2013

Aceite a 9 de março de 2015

### Palavras-chave:

Reprodutibilidade

Teste de campo

Criança

Adolescente

## R E S U M O

**Objetivo:** O objetivo do estudo foi verificar a confiabilidade do teste de corrida/caminhada de 9 minutos em crianças e adolescentes de 7-12 anos de idade.

**Método:** Participaram do presente estudo 95 escolares de ambos os sexos (54 meninos e 41 meninas) do município de Londrina-PR, com idade média de  $9.4 \pm 1.8$  anos; massa corporal  $31.7 \pm 13.1$  kg e estatura  $137.0 \pm 15.2$  cm. Os sujeitos realizaram o teste de corrida/caminhada de 9 minutos em uma pista de atletismo em 2 momentos, sendo separados por uma semana de intervalo entre cada medida, para avaliação do teste-reteste. A confiabilidade foi testada por testes estatísticos para esta finalidade considerando uma significância de 5%.

**Resultados:** O teste de 9 minutos apresentou um erro técnico de medida absoluto = 110.4 m e relativo = 9.1% e uma variação de 12.5%. Uma confiabilidade de  $r=0.85$  e a maioria dos resultados apresentaram-se dentro dos limites de concordância 95% (magnitude de  $21.1 \pm 304.9$  m).

**Conclusão:** Com base nos resultados observou-se que o teste de corrida/caminhada de 9 minutos apresenta resultados aceitáveis em relação aos cálculos realizados para verificação da confiabilidade em escolares de 7-12 anos de idade.

© 2013 Consejería de Educación, Cultura y Deporte de la Junta de Andalucía. Publicado por Elsevier España, S.L.U. Este é um artigo Open Access sob a licença de CC BY-NC-ND

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## La fiabilidad del test de correr-caminar de 9 minutos en niños y adolescentes de 7-12 años de edad

## R E S U M E N

**Objetivo:** El objetivo del estudio fue analizar la fiabilidad del test de correr/caminar durante 9 minutos en niños y adolescentes de 7-12 años de edad.

**Método:** El estudio incluyó a 95 escolares de ambos sexos (54 niños y 41 niñas) de Londrina-PR con una edad media de  $9.4 \pm 1.8$  años, peso corporal de  $31.7 \pm 13.1$  kg y altura de  $137.0 \pm 15.2$  cm. Los sujetos realizaron el test de caminata de 9 minutos en una pista de atletismo en 2 momentos separados por una semana para evaluación del test-retest. La fiabilidad fue valorada por tests estadísticos con un nivel de significación del 5%.

**Resultados:** El test de 9 minutos presentó un error técnico de medida absoluto de 110.4 m, un error relativo de 9.1% y una variación del 12.5%. La mayoría de los resultados se encontraron dentro de los límites de confianza del 95% ( $21.1 \pm 304.9$  m).

### Palabras clave:

Reproducibilidad

Pruebas de campo

Niños

Adolescentes

\* Autor para correspondência.

Correio eletrónico: [anacarolinapaludo@usp.br](mailto:anacarolinapaludo@usp.br) (A.C. Paludo).

**Conclusión:** Con base en los resultados, se observa que el test de correr-caminar de 9 minutos en escolares de 7-12 años de edad muestra resultados aceptables en relación a los cálculos realizados para determinar la fiabilidad.

© 2013 Consejería de Educación, Cultura y Deporte de la Junta de Andalucía. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Confiability of the run/walk 9 minutes test in children and adolescents from 7 to 12 years old

A B S T R A C T

**Keywords:**  
Reproducibility  
Field test  
Children  
Adolescents

**Objective:** The purpose of this study was to verify the confiability of the test run/walk 9 minutes in children and adolescents from seven to 12 years old.

**Method:** The study included 95 schoolchildren of both sexes (54 boys and 41 girls) from Londrina- PR with a mean age of  $9.44 \pm 1.80$  years old, body weight  $31.70 \pm 13.10$  kg and height  $137 \pm 15.2$  cm. Subjects performed the test run/walk 9 minutes on a running track two times being separated by a week between each measure for evaluation of the test–retest. Confiability was tested by statistical testes, with a significance of 5%.

**Results:** The 9 minutes test presented an absolute error of 110.4 m, a relative error of 9.1% and a coefficient of variation of 12.5%. Most of the results were within the 95% limits of agreement (magnitude of  $21.1 \pm 304.9$  m).

**Conclusion:** The run/walk 9 minutes test shows acceptable results in relation to the calculation performed to determine the reproducibility in schoolchildren from 7 to 12 years old.

© 2013 Consejería de Educación, Cultura y Deporte de la Junta de Andalucía. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Introdução

A aptidão cardiorrespiratória (ACR), componente da aptidão física tanto voltada para a saúde quanto para o desempenho atlético, vem sendo amplamente apontada na literatura como um importante aspecto de saúde, visto que níveis adequados desse componente estão associados a baixo risco de doenças cardiovasculares e mortalidade, na população adulta<sup>1</sup>. Em populações pediátricas o mesmo benefício pode ser observado, sendo que índices satisfatórios da ACR estão inversamente relacionados com fatores de risco, tais como excesso de adiposidade abdominal e total, bem como perfil cardiovascular e metabólico desfavoráveis<sup>2-4</sup>.

Assim, a avaliação da ACR torna-se uma importante ferramenta para quantificar, identificar e prevenir possíveis fatores de risco em diferentes populações, como estratégia em âmbito de saúde pública. Neste sentido, a mensuração do consumo máximo de oxigênio ( $VO_{2\text{máx}}$ ) por meio da avaliação da potência aeróbia máxima, vem sendo reconhecido como um dos melhores índices para determinação da ACR em adultos e jovens. O  $VO_{2\text{máx}}$  pode ser avaliado pela análise direta, envolvendo métodos laboratoriais ou por análise indireta, caracterizado pelos testes motores de campo no qual envolvem corridas e/ou caminhadas de diferentes distâncias<sup>5</sup>.

Os testes de campo têm sido uma das alternativas mais utilizadas para avaliação da ACR na população de crianças e adolescentes, por apresentarem importantes vantagens como: o baixo custo operacional; a facilidade de aplicação dos protocolos de teste; avaliação de um grande número de sujeitos simultaneamente; maior acessibilidade aos locais de teste, consequentemente permitindo a realização do teste fora do ambiente laboratorial, bem como utilizar-se dos resultados obtidos no teste para estimar a ACR<sup>6</sup>.

Contudo, apesar da viabilidade dos testes de campo, é imprescindível que seus respectivos protocolos forneçam uma boa estimativa do que se propõem a medir, apresentando boa validade e reprodutibilidade. A validade refere-se à habilidade do teste de medir aquilo que se propõe a medir, enquanto a reprodutibilidade reflete a consistência das medidas, no qual são realizadas

repetidas avaliações em um mesmo indivíduo sobre as mesmas condições<sup>7</sup>.

A reprodutibilidade de testes de campo ainda é pouco explorada, principalmente quando levamos em conta os testes existentes na literatura para a população de crianças e adolescentes. No Brasil, um dos testes mais utilizados para estimativa da ACR em jovens é o teste de corrida/caminhada de 9 minutos (9 min). Ele vem sendo sugerido como alternativa para a avaliação da ACR em baterias de testes motores nacionais<sup>8</sup>, bem como internacionais<sup>9</sup>. Além disso, fornece pontos de corte para triagem de fatores de risco para doenças cardiovasculares na infância<sup>10</sup>.

Diante desse fato e considerando os testes de campo para estimativa da ACR como ferramentas importantes e acessíveis de diagnóstico e prevenção de fatores de risco à saúde nos jovens, destaca-se a necessidade da investigação dos critérios relacionados à reprodutibilidade destes protocolos. Nesse sentido, o objetivo do presente estudo foi verificar a reprodutibilidade do teste de corrida/caminhada de 9 min em crianças e adolescentes de 7-12 anos de idade, por sexo.

## Método

### Amostra

O presente estudo faz parte do banco de dados do projeto intitulado «Comparação da potência aeróbia estimada mediante a aplicação de 3 diferentes testes de campo em adolescentes», vinculado à Universidade Estadual de Londrina, do qual fizeram parte 288 escolares matriculados entre a 2.<sup>a</sup> e 6.<sup>a</sup> séries do ensino fundamental, de 2 instituições da rede pública de ensino do município de Londrina-PR, selecionadas por conveniência.

Do total de escolares participantes, foram sorteadas algumas turmas para realizar a repetição do teste. Foram incluídos nas análises aproximadamente 33% dos escolares que compuseram a amostra total, somando 95 sujeitos (54 meninos e 41 meninas) com as medidas repetidas do teste de 9 min.

Como critérios de inclusão foram considerados os seguintes aspectos: pertencer às séries preestabelecidas e estar regularmente matriculados nas instituições de ensino selecionadas no ano letivo de 2009 e no primeiro semestre de 2011. Como critérios de exclusão foram estabelecidas as seguintes condições: recusa em participar do estudo, não autorização dos pais ou responsáveis, apresentar algum problema físico que impedisse o indivíduo a realizar os testes motores e a ausência às aulas no dia da coleta dos dados.

Todos os alunos envolvidos na pesquisa e seus respectivos responsáveis foram informados previamente quanto aos objetivos do estudo e receberam esclarecimento sobre os procedimentos adotados. Os responsáveis assinaram um Termo de Consentimento Livre e Esclarecido e o projeto foi aprovado pelo Comitê de Ética da Universidade Estadual de Londrina.

### Antropometria

A massa corporal dos sujeitos foi obtida em uma balança digital com precisão de 0.1 kg (modelo PS 180A; Urano, Porto Alegre, Brasil) e a estatura foi determinada por um estadiômetro de madeira com precisão de 0.1 cm, de acordo com os procedimentos descritos por Gordon et al.<sup>11</sup>. Todos os indivíduos foram medidos e pesados com roupas leves e descalços. Com base nessas informações, calculou-se o índice de massa corporal ( $IMC = kg/m^2$ ).

### Testes de 9 minutos

O teste de campo de 9 min foi realizado em uma pista de atletismo oficial, seguindo as recomendações de Cooper<sup>12</sup>, no qual os sujeitos foram orientados a caminhar e/ou correr a maior distância possível no tempo de 9 min. Controlou-se a distância através do número de voltas completas na pista de 400 m, somando-se os metros adicionais.

Os sujeitos realizaram o mesmo protocolo do teste em 2 momentos, nas mesmas condições (período do dia), sendo separados por uma semana de intervalo entre as medidas.

### Tratamento estatístico

Para verificar a distribuição dos dados foi utilizado o teste de *Shapiro Wilk*, através do qual se detectou que as variáveis antropométricas não apresentaram distribuição normal (descritas em mediana e intervalo interquartil). Assim, a comparação destas características por sexo foi estabelecida pelo teste *U* de *Mann-Whitney*. Por outro lado, os valores obtidos nos testes de 9 min foram analisados por meio dos testes paramétricos.

Para o cálculo da reprodutibilidade foi utilizado: erro técnico de medida (ETM) expresso na sua forma relativa e absoluta. O ETM absoluto foi calculado pela raiz quadrada da soma das diferenças entre as 2 aplicações, ao quadrado, dividida por 2 vezes o número de pares<sup>7</sup>.

$$ETM(\text{absoluto}) = \sqrt{\frac{\sum d_i^2}{2n}}$$

No qual:  $\sum d$  = somatório dos desvios elevado ao quadrado.  $n$  = número de voluntários medidos (testados).  $i$  = quantas forem as diferenças.

Enquanto o ETM relativo foi estabelecido de acordo com Peder-son e Gore<sup>13</sup>, mediante a divisão do ETM absoluto pelo valor médio da variável multiplicado por 100.

$$ETM(\text{relativo}) = \frac{ETM(\text{absoluto})}{VMV} * 100$$

No qual: ETM (absoluto) = valor do ETM absoluto. VMV = valor médio da variável.

**Tabela 1**

Características descritivas da amostra apresentadas em mediana (intervalo interquartil)

| Variáveis                | Meninos (n = 54) | Meninas (n = 41) | Total (n = 95) |
|--------------------------|------------------|------------------|----------------|
| Idade (anos)             | 9.5 (1.6)        | 9.0 (1.8)        | 9.4 (1.8)      |
| Massa corporal (kg)      | 31.6 (9.2)       | 32.8 (16.7)      | 31.7 (13.1)    |
| Estatura (cm)            | 137.0 (15.2)     | 135.5 (14.7)     | 136.5 (18.0)   |
| IMC (kg/m <sup>2</sup> ) | 16.8 (3.0)       | 17.8 (6.0)       | 17.2 (4.3)     |

IMC: índice de massa corporal.

Ainda foi utilizado o coeficiente de variação (CV) expresso pela divisão do erro padrão de estimativa (EPE) pela média da segunda medida (MSM) do teste de campo, multiplicado por 100.

$$CV = \frac{EPE}{MSM} * 100$$

Adicionalmente, o coeficiente de correlação intraclass (CCI) foi utilizado para verificar a consistência das medidas, a partir de informações da análise de variância de medidas repetidas. O teste «t» de *Student* pareado possibilitou a comparação entre as médias das distâncias percorridas no teste de 9 min no momento 1 e 2 e, por fim, a plotagem de Bland e Altman<sup>14</sup> foi utilizada para análise individualizada dos limites de concordância (95%) entre as medidas repetidas. Utilizou-se a significância estatística de 5% para todas as análises.

### Resultados

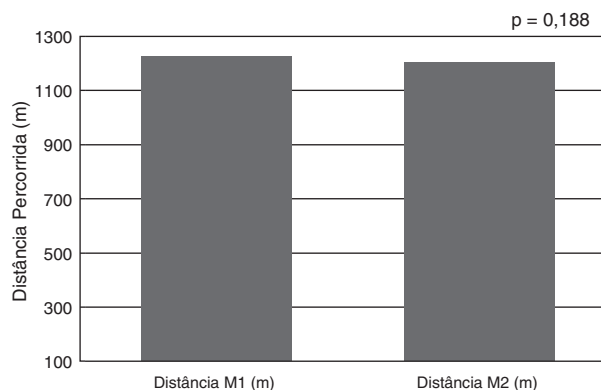
As características descritivas da amostra são apresentadas na [tabela 1](#). O teste *U* de *Mann-Whitney* não indicou nenhuma diferença significativa nas variáveis descritivas entre os sexos ( $p > 0.05$ ).

A distância média percorrida pelos sujeitos no primeiro momento em que o teste foi aplicado foi de  $1224.0 \pm 210.25$  m ( $1253.0 \pm 241.0$  m para os meninos e  $1185.6 \pm 156.2$  m para as meninas) e na réplica do teste a média foi de  $1202.0 \pm 221.28$  m ( $1225.0 \pm 231.2$  m para os meninos e  $1173.4 \pm 206.5$  m para as meninas), não ocorrendo diferença estatisticamente significativa ( $p = 0.188$ ), como apresentado na [figura 1](#).

O cálculo dos ETM (absoluto e relativo), CV e CCI, referentes à análise da reprodutibilidade do teste de 9 min, são apresentados na [tabela 2](#).

Os dados de concordância da amostra total ( $n = 95$ ) podem ser visualizados na [figura 2](#).

A disposição dos pontos na [figura 2](#) demonstra que grande parte dos valores estão localizados entre os limites de concordância  $\pm 1.96$  desvio padrão, mas com 5 indivíduos fora destes limites (outliers). Os limites de concordância apresentam-se na magnitude



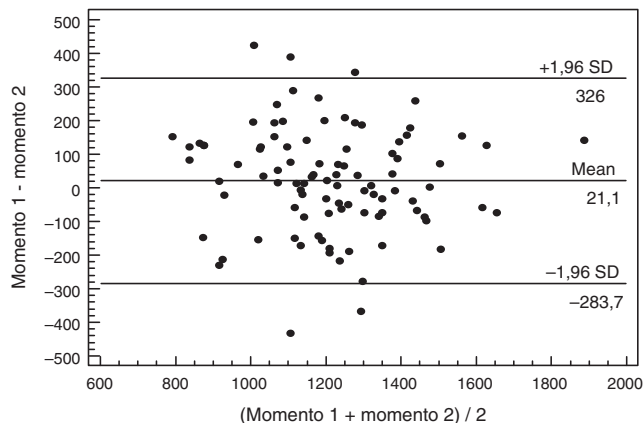
**Figura 1.** Distância média (metros) percorrida no teste de 9 min no momento um e momento 2.

**Tabela 2**

Cálculo da reprodutibilidade do teste de 9 min

| Testes           | Meninos (n=54) | Meninas (n=41) | Total (n=95) |
|------------------|----------------|----------------|--------------|
| ETM absoluto (m) | 118,6          | 98,6           | 110,4        |
| ETM relativo (%) | 9,6            | 8,4            | 9,1          |
| CV (%)           | 12,5           | 12,1           | 12,5         |
| CCI              | 0,9            | 0,8            | 0,8          |

CCI: coeficiente de correlação intraclasse; CV: coeficiente de variação; ETM: erro técnico de medida.

**Figura 2.** Concordância entre o momento um e momento 2 do teste de corrida/caminhada 9 min.

de  $21.1 \pm 304.9$  m, ou seja, em 95% dos casos um mesmo jovem que realiza o teste de 9 min em 2 momentos distintos no tempo pode apresentar uma variação de aproximadamente 305 m, na distância total percorrida. Além disso, de maneira geral, os valores de distância avaliados pelo teste de 9 min no momento 1 foram superiores aos valores encontrados no momento 2, no entanto, o viés de medida entre os momentos (21.1 m) não foi considerado diferente de 0 ( $p=0.188$ ), o que seria esperado para uma boa medida de concordância.

## Discussão

O presente estudo apresentou como principais resultados que, os valores de reprodutibilidade podem ser considerados aceitáveis para o teste de 9 min, na amostra estudada.

Ao contrário das medidas antropométricas, nas quais existem pontos de corte que fornecem informações sobre os níveis das medidas de teste-reteste<sup>15</sup>, os testes motores e em particular os testes para avaliação da ACR, ainda não apresentam classificações e valores de referência para as medidas da reprodutibilidade.

Artigos de revisão encontrados na literatura<sup>6,7,16</sup> enfatizam a importância de verificar a reprodutibilidade em testes que visam o desempenho, principalmente em testes de campo, no qual o erro aparenta ser maior<sup>7,16</sup>, uma vez que as medidas de reprodutibilidade indicam a variação biológica e a técnica dos protocolos<sup>17</sup>, permitindo verificar a extensão do erro e viés dos testes. Neste sentido, a reprodutibilidade de um teste pode ser avaliada de forma relativa e absoluta. A forma relativa refere-se ao grau com que cada indivíduo mantém sua posição na mesma amostra sobre a medida repetida e a forma absoluta refere-se ao grau em que cada medida repetida varia para os indivíduos<sup>18</sup>.

Buscando avaliar a extensão de ambas as formas de expressão, o presente estudo analisou a reprodutibilidade relativa pelo teste estatístico de CCI e absoluta com os testes ETM, CV e análise de concordância de Bland e Altman, recomendados pela literatura para essa finalidade<sup>19</sup>.

Os resultados para a amostra total apontam para um ETM relativo de 9.1% e um CV de 12.5%, este último um pouco acima dos 10% recomendados pela literatura<sup>16</sup>. Seu CCI foi de  $r=0.85$ , que significa uma reprodutibilidade aceitável, conforme a classificação para os coeficientes de correlação propostos por Vincent<sup>20</sup> (alto:  $\geq 0.90$ ; moderado: 0.70-0.89; baixo: 0.50-0.70). Vale ressaltar que para os meninos, isoladamente, a classificação do CCI pode considerar a reprodutibilidade como alta ( $r=0.90$ ) comparado as meninas ( $r=0.80$ ).

A plotagem de Bland e Altman, que permite verificar limites de concordância entre as 2 medidas, pode ser entendida na análise de reprodutibilidade como o intervalo de tolerância da consistência entre os 2 testes<sup>21</sup>. Sendo assim, o esperado é que ao realizar 2 vezes o teste de 9 min, com os mesmos sujeitos em condições semelhantes, as diferenças entre o momento 1 e momento 2 sejam o mais próximas possível do valor 0 e o intervalo de confiança com os limites extremos próximos dos valores referentes às diferenças médias. Para o presente estudo, conforme visualizado no diagrama de dispersão (fig. 2), a diferença média apresentada entre os 2 momentos foi de 21.1 m. Em relação ao intervalo de confiança, os limites superior e inferior foram de 326 e -283.7 m, respectivamente.

Recentemente, Artero et al.<sup>22</sup> reuniram informações em uma revisão sistemática acerca da reprodutibilidade dos testes de campo realizados em crianças e adolescentes para estimativa da ACR. Dos estudos encontrados na literatura, os testes que foram analisados a reprodutibilidade foram o teste de uma milha, ½ milha, *shuttle-run* de 20 m, caminhada de 6 min, corrida de 5 min. O teste com melhor resultado de reprodutibilidade, classificado pelos autores como «forte», foi o teste de *shuttle-run* de 20 m e o teste de uma milha com «moderada» evidência.

Os resultados encontrados pelo presente estudo para o teste de 9 min são semelhantes aos resultados apresentados pelos testes analisados na revisão de Artero et al.<sup>22</sup>. A maioria dos trabalhos utilizaram as análises estatísticas de CCI e os limites de concordância apresentados na plotagem de Bland e Altman com o objetivo de analisar a reprodutibilidade dos testes. O CCI de  $r=0.85$  para a amostra total do presente estudo encontra-se dentro dos valores de consistência apresentados nos resultados apontados pelos testes com a mesma finalidade, conforme resumizam Artero et al.<sup>22</sup>.

Ao comparar os valores obtidos para o teste de 9 min no presente trabalho, com o teste de corrida/caminhada de 6 min do estudo de Li et al.<sup>23</sup>, no qual apresentam as mesmas características (teste com tempo fixo), seus valores encontraram-se inferiores. O CCI do teste de corrida/caminhada de 6 min foi maior (0.94) e a média das diferenças e os limites de concordância foram menores (15 m; 65 m e -35 m, respectivamente) em relação ao teste de 9 min.

Tal comportamento parece ser explicado pelo fato da reprodutibilidade de testes com duração superior a 1 min, como no caso dos testes para verificação da ACR, apresentarem baixos valores de reprodutibilidade devido a variações no ambiente de realização dos testes e no *status* do avaliado incluindo seu estado nutricional, fadiga e motivação<sup>24</sup>. Enfim, parece haver uma relação inversa entre tempo total de teste e variáveis relacionadas à sua validade, entre elas a reprodutibilidade.

Por fim, pode-se apontar como possíveis limitações encontradas na realização do presente estudo o controle das condições climáticas. Embora a temperatura e umidade relativa do ar não tenham sido exatamente os mesmos, tentou-se controlar os dias com temperaturas semelhantes, optando pela realização dos testes no mesmo período (tarde).

Por outro lado, um dos pontos positivos do estudo foi a utilização de todas as ferramentas estatísticas de verificação da reprodutibilidade indicadas na literatura, permitindo uma análise mais consistente dos resultados, tanto considerando valores médios quanto individuais. Adicionalmente, o presente estudo procurou manter os mesmos avaliadores no teste-reteste, favorecendo a

reprodução das mesmas condições de orientação e motivação repassada aos sujeitos.

Em resumo, o teste de corrida/caminhada de 9 min apresentou valores aceitáveis de reprodutibilidade, tanto em termos absolutos quanto em termos relativos, para a amostra de crianças e adolescentes estudada.

### Financiamiento

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelas bolsas de produtividade em pesquisa (E.R.V.R e E.S.C.), a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de Doutorado outorgada (A.C.P e M.B.B.) e a Fundação de Apoio ao Desenvolvimento Científico e Tecnológico de Paraná (FAADCT/PR) pelo financiamento ao projeto.

### Conflito de interesses

Os autores declaram não haver conflito de interesses.

### Referências

1. Kodama S, Saito K, Tanaka S, Maki M, Yachi Y, Asumi M, et al. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: A meta-analysis. *JAMA*. 2009;301:2024–35.
2. Eisenmann JC, Welk GJ, Ihmels M, Dollman J. Fatness, fitness, and a cardiovascular disease risk factor in children and adolescents. *Med Sci Sports Exerc*. 2007;39:1251–6.
3. Anderssen SA, Cooper AR, Riddoch C, Sardinha LB, Harro M, Brage S, et al. Low cardiorespiratory fitness is a strong predictor for clustering of cardiovascular disease risk factor in children independent of country, age and sex. *Eur J Cardiovasc Prev Rehabil*. 2007;14:526–31.
4. Ortega FB, Ruiz JR, Castillo MJ, Sjostrom M. Physical fitness in childhood and adolescence: A powerful marker of health. *Int J Obes*. 2008;32:1–11.
5. Guedes DP, Guedes JERP. Avaliação de aspectos funcionais: Sistema de mobilização em energia. In: Guedes DP, Guedes JERP, editors. *Manual prático para avaliação em educação física*. 1ª ed São Paulo: Manole; 2006. p. 346–415.
6. Castro-Piñero J, Artero EG, España-Romero V, Ortega FB, Sjöström M, Suni J, et al. Criterion-related validity of field-based fitness tests in youth: A systematic review. *Br J Sports Med*. 2010;44:934–43.
7. Currell K, Jeukendrup AE. Validity, reliability and sensitivity of measures of sporting performance. *Sports Med*. 2008;38:297–316.
8. Gaya ACA. Projeto Esporte Brasil: PROESP-BR. Manual de Aplicação de Medidas e Testes, Normas e Critérios de Avaliação. Porto Alegre-RS, 2009.
9. American Alliance for Health, Physical Education, Recreation, and Dance (AAHPERD). *Physical Best: The American Alliance physical fitness education and assessment program*. Reston, VA: AAHPERD; 1980.
10. Bergmann GG, Gaya ACA, Halpen R, Bergmann MLA, Rech RR, Constanzi CB, et al. Pontos de corte para a aptidão cardiorrespiratória e a triagem de fatores de risco para doenças cardiovasculares na infância. *Rev Bras Med Esporte*. 2010;16:339–43.
11. Gordon CC, Chumlea WC, Roche AF. Stature, recumbent length, and weight. In: Lohman TG, Roche AF, Martorelli R, editors. *Anthropometric standardizing reference manual*. Champaign, Illinois: Human Kinetics Books; 1988. p. 3–8.
12. Cooper KH. A means of assessing maximal oxygen uptake. *J Am Med Assoc*. 1968;203:135–8.
13. Pederson D, Gore C. Erros de medição em antropometria. In: Norton K, Olds T, editors. *Antropométrica*. Porto Alegre: Artmed; 2005. p. 71–86.
14. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307–10.
15. Norton K, Olds T. *Antropométrica*. Argentina: Biosystem; 2000.
16. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sport medicine. *Sports Med*. 1998;26:217–38.
17. Bagger M, Petersen PH, Pedersen PK. Biological variation in variables associated with exercise training. *Int J Sports Med*. 2003;24:433–40.
18. Baumgartner TA. Norm-referenced measurement: reliability. In: Safran MJ, Wood TM. (eds.). *Measurement Concepts in Physical Education and Exercise Science*. Champaign, Illinois; 1989. p. 45–72.
19. Bruton A, Conway JH, Holgate ST. Reliability: What is it, and how is it measured. *Physiotherapy*. 2000;86:94–9.
20. Vincent W. *Statistics in kinesiology*. 3rd ed Champaign (IL): Human Kinetics; 2005.
21. Chatburn RL. Evaluating of instrument error and method agreement. *AANA J*. 1996;64:261–8.
22. Artero EG, España-Romero V, Castro-Piñero J, Ortega FB, Suni J, Castillo-Garzon MJ, et al. Reliability of field-based fitness tests in youth. *Int J Sports Med*. 2011;32:159–69.
23. Li AM, Yin J, Yu CC, Tsang T, So HK, Chan D, et al. The six-minute walk test in healthy children: Reliability and validity. *Eur Respir J*. 2005;25:1057–60.
24. Hopkins WG, Schabert EJ, Hawley JA. Reliability of power in physical performance tests. *Sports Med*. 2001;31:211–34.