



# CIÊNCIAS DO ESPORTE

[www.rbceonline.org.br](http://www.rbceonline.org.br)



## ARTIGO ORIGINAL

# Reprodutibilidade do desempenho em provas de corrida de 5 e 10 km em pista de atletismo



Fabiana Andrade Machado<sup>a,b,\*</sup>, Ana Claudia Pelissari Kravchychyn<sup>b</sup>,  
Cecília Segabinazi Peserico<sup>b</sup>, Danilo Fernandes da Silva<sup>b</sup> e Paulo Victor Mezzaroba<sup>b</sup>

<sup>a</sup> Departamento de Educação Física, Centro de Ciências da Saúde, Universidade Estadual de Maringá (UEM), Maringá, PR, Brasil

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

Recebido em 3 de março de 2012; aceito em 2 de setembro de 2012

Disponível na Internet em 3 de junho de 2015

### PALAVRAS-CHAVE

Coefficiente de variação;  
Performance;  
Tamanho da amostra;  
Teste-reteste

**Resumo** Examinou-se a reprodutibilidade do desempenho em provas de corrida de 5 e 10 km em pista de atletismo. Trinta e um corredores recreacionistas, do gênero masculino, fizeram duas provas de 5 e duas de 10 km em uma pista oficial de atletismo. Não houve diferença significativa entre os testes-retestes para ambas as provas com o uso do teste *t* pareado ( $p > 0,05$ ). Os coeficientes de correlação intraclasse e de variação foram de 0,94/3,44% (5 km) e de 0,97/2,43% (10 km), respectivamente. A reprodutibilidade de ambas as provas em pista de atletismo foi semelhante às reprodutibilidades reportadas na literatura para testes em laboratório. Além disso, a prova de 10 km se mostrou mais reprodutível do que a prova de 5 km para a amostra estudada e necessitou de aproximadamente metade da amostra necessária para se detectarem aumentos significativos durante teste-reteste em provas de 5 km.

© 2015 Colégio Brasileiro de Ciências do Esporte. Publicado por Elsevier Editora Ltda. Todos os direitos reservados.

### KEYWORDS

Coefficient of variation;  
Performance;  
Sample size;  
Test-retest

### Reproducibility of the performance in running events of 5 and 10 km in running track

**Abstract** We examined the reproducibility of performance in running events of 5 and 10 km in running track. Thirty-one recreational, male, runners performed two tests of 5 and two tests of 10 km in an official running track. There was no significant difference in time between the test-retest for both tests using the paired *t* test ( $P > 0.05$ ). The intraclass coefficient of correlation and the coefficient of variation were 0.94/3.44% (5 km) and 0.97/2.43% (10 km), respectively. The reproducibility of both running events in running track was similar to that reported in the literature for laboratory tests. In addition, the 10 km test was more reproducible than the

\* Autor para correspondência.

E-mail: [famachado.uem@hotmail.com](mailto:famachado.uem@hotmail.com) (F.A. Machado).

**PALABRAS CLAVE**

Coefficiente de variación; Rendimiento; Tamaño de la muestra; Test-retest

5 km test for this studied sample, requiring approximately half the sample needed to detect significant increases during test-retest in running events of 5 km.

© 2015 Colégio Brasileiro de Ciências do Esporte. Published by Elsevier Editora Ltda. All rights reserved.

**Reproducibilidad del desempeño en pruebas de carrera de 5 y 10 km en pista de atletismo**

**Resumen** Hemos examinado la reproducibilidad del desempeño en pruebas de carrera de 5 y 10 km en pista de atletismo. Treinta y un corredores recreativos, género masculino, realizaron dos pruebas de 5 y dos de 10 km en una pista oficial de atletismo. No hubo diferencia significativa en el tiempo test-retest entre ambas pruebas mediante la prueba de t para muestras apareadas ( $P > 0,05$ ). Los coeficientes de correlación intraclase y de variación fueron de 0,94/3,44% (5 km) y 0,97/2,43% (10 km), respectivamente. Se encontró que la reproducibilidad de las dos pruebas de carrera en pista de atletismo fue similar a la descrita en la literatura para pruebas de laboratorio. Además, la prueba de 10 km fue más reproducible que la prueba de 5 km para la muestra estudiada, lo que requiere aproximadamente la mitad de la muestra necesaria para detectar un aumento significativo en test-retest en pruebas de 5 km.

© 2015 Colégio Brasileiro de Ciências do Esporte. Publicado por Elsevier Editora Ltda. Todos los derechos reservados.

**Introdução**

O desempenho durante o exercício físico prolongado vem sendo amplamente usado para se avaliarem os efeitos do treinamento e de recursos ergogênicos sobre a corrida e o ciclismo (Beelen et al., 2009; Carter et al., 2004; Chen et al., 2008; Lindsey et al., 1996; Madsen et al., 1996; Rollo e Williams, 2010; Rollo et al., 2011; Whitham e Mckinney, 2007). Para tanto, esse desempenho deve ser altamente reprodutível a fim de que se detectem pequenas, mas significativas, melhorias no desempenho (Atkinson e Nevill, 1998; Hopkins et al., 1999; Hopkins, 2000; Hopkins e Hewson, 2001; Hopkins et al., 2001).

Embora diversos estudos tenham avaliado a reprodutibilidade do desempenho de *endurance* durante o ciclismo (Bishop, 1997; Hickey et al., 1992; Jensen e Johansen, 1998; Jeukendrup et al., 1996; Palmer et al., 1996; Schabort et al., 1998a; Smith et al., 2001), são poucos os estudos que examinaram essa reprodutibilidade durante a corrida (Doyle e Martinez, 1998; Schabort et al., 1998b; Rollo et al., 2008; Russel et al., 2004). Além disso, esses poucos estudos durante a corrida foram feitos em laboratórios, carece a literatura de estudos em pistas de atletismo, os quais seriam úteis a treinadores e pesquisadores que pretendam analisar os efeitos do treinamento e de recursos ergogênicos diretamente em provas de longa distância.

Schabort et al. (1998b) determinaram, em oito corredores treinados, a reprodutibilidade da distância percorrida durante 60 minutos em uma esteira ergométrica motorizada em que os participantes ajustavam a velocidade manualmente. O coeficiente de variação (CV) foi de 2,7% (IC 95% = 1,8-4,0%) e a reprodutibilidade teste-reteste expressa pelo coeficiente de correlação intraclasse (CCI) foi de 0,90 (IC 95% = 0,72-0,98). Rollo et al. (2008) determinaram essa

mesma reprodutibilidade em 10 corredores treinados. Entretanto, usaram uma esteira ergométrica em que a velocidade de corrida foi controlada automaticamente, sem a necessidade de ajustes manuais, e encontraram um CV bem menor (1,4%). Doyle e Martinez (1998) examinaram em 10 corredores treinados a reprodutibilidade do tempo para se percorrer uma distância média de 5,05 km após uma corrida de 90 minutos em uma esteira ergométrica a uma intensidade de 70% do consumo máximo de oxigênio ( $VO_2\max$ ) e relataram um elevado CV (4,4%). Já Russell et al. (2004) avaliaram em oito corredores de *endurance* a reprodutibilidade do tempo para se percorrer 10 km após uma corrida de 90 minutos a uma intensidade de 65% do  $VO_2\max$  e relataram um CV bem baixo (1,0%).

Outro aspecto relacionado à reprodutibilidade diz respeito ao tamanho da amostra necessária para se detectarem mudanças significativas na variável estudada devido ao treinamento ou uso de recursos ergogênicos (Hopkins et al., 1999; Hopkins, 2000). Sabe-se, por exemplo, que o treinamento de *endurance* proporciona uma melhoria mais acentuada em indivíduos não treinados do que em corredores de elite, mas pequenos aumentos no desempenho de atletas de elite podem levá-lo a vencer ou perder uma prova (Schabort et al., 1998b). Dessa forma, dependendo do aumento esperado no desempenho e considerando o tamanho da amostra determinada, é possível que a reprodutibilidade de uma prova de 5 ou 10 km se mostre suficiente para avaliar os efeitos do treinamento em indivíduos não treinados, mas inadequada para a avaliação de corredores de elite.

Diante disso, com vistas a avaliar a reprodutibilidade do desempenho de *endurance* em um ambiente mais próximo ao que ocorrem as competições, este estudo teve como objetivos: 1) examinar a reprodutibilidade do desempenho

em provas de corrida de 5 e 10 km em pista de atletismo e 2) determinar o tamanho da amostra necessário para se verificarem aumentos significativos no desempenho durante teste-reteste em provas de 5 e 10 km.

## Material e métodos

### Sujeitos

Participaram deste estudo 31 corredores recreacionistas, do gênero masculino, aparentemente saudáveis, integrantes de grupos locais de corrida. Os participantes tinham tempos em provas de corrida de 10 km entre 35 e 60 minutos, com uma velocidade média entre 10 e 17 km/h (~44-75% do recorde mundial). Os corredores foram selecionados de acordo com os seguintes critérios: experiência mínima na modalidade (dois anos) e regularidade de treinamento (pelo menos duas vezes por semana). As características dos participantes (média ± DP) foram: idade 39,4 ± 12,8 anos; estatura 173,5 ± 7,5 cm; massa corporal 68,2 ± 8,5 kg; e índice de massa corporal (IMC) 22,6 ± 2,1 kg/m<sup>2</sup>. Os participantes treinavam com uma frequência de 4,9 ± 1,5 sessões por semana e percorriam uma distância semanal de 58,5 ± 31,7 km. Todos os participantes leram e assinaram o termo de consentimento livre e esclarecido antes do início dos testes. O protocolo usado neste estudo foi previamente aprovado pelo Comitê de Ética local (# 719/2010).

### Desempenho em provas de 5 e 10 km

Todos os participantes deste estudo fizeram duas provas de 5 e duas de 10 km em uma pista oficial de atletismo de 400 m ao longo de duas semanas. Os participantes foram instruídos a comparecer ao local das provas bem hidratados, com roupas leves e confortáveis e a última refeição feita com pelo menos duas horas de antecedência do início das provas. Também foram instruídos a se abster de bebidas cafeinadas ou alcoólicas e de exercícios extenuantes nas 24 horas precedentes aos testes e a consumir a mesma dieta e a manter o mesmo regime de treinamento nas 48 horas precedentes aos testes. As provas de 5 km foram feitas na primeira semana, com intervalo mínimo de 48 horas entre o teste-reteste, e as provas de 10 km foram feitas na segunda semana, similarmente com intervalo mínimo de 48 horas entre o teste-reteste. As provas iniciaram às 18 horas, precedidas de um aquecimento livre de 10 minutos. Em um dos pontos da pista de atletismo foi montada uma estação para hidratação em que havia água mineral para os participantes. Todos os participantes foram incentivados e encorajados ao longo das provas para obter o melhor desempenho possível. O tempo de cada prova foi registrado com o auxílio de cronômetros e usado como variável de desempenho para as análises de reprodutibilidade.

### Tamanho da amostra

A determinação do tamanho da amostra necessário para se verificarem aumentos significativos ( $n_{sign}$ ) no desempenho durante teste-reteste em provas de 5 e 10 km foi feita com a seguinte equação, que somente é válida para teste-reteste

unilateral, sem grupo controle, para um nível de significância de 5% e um poder estatístico de 80% (Hopkins, 2000):

$$n_{sign} = 2 \times (t_{0,975,n-1} \times CV/D)^2$$

em que  $t_{0,975,n-1}$  é obtido da tabela  $t$  de Student para um nível de confiança unicaudal de 97,5% e  $n-1$  graus de liberdade;  $CV$  é o coeficiente de variação da prova (5 ou 10 km) e  $D$  é a porcentagem de aumento esperado no desempenho, representado aqui pela melhoria esperada no tempo das provas de 5 ou 10 km devido ao treinamento ou a recursos ergogênicos.

### Análise estatística

Os dados são apresentados como média ± desvio padrão (DP). As análises foram feitas com o Statistical Package for the Social Sciences versão 17.0 Trial (SPSS Inc. USA). A normalidade dos dados foi verificada por meio do teste de Shapiro-Wilk. O teste de Levene verificou a homogeneidade das variâncias. As variáveis foram comparadas por meio do teste  $t$  de Student bicaudal pareado, o qual foi usado na análise de erro sistemático. Os limites de concordância propostos por Bland e Altman (1986) para um intervalo de confiança de 95% foram usados na análise da reprodutibilidade do desempenho das provas de 5 e 10 km. O CCI de consistência interna e o coeficiente de correlação de Pearson ( $r$ ) examinaram a reprodutibilidade relativa, em que o CCI foi calculado por meio do modelo misto para uma única medida. Além disso, a reprodutibilidade absoluta entre o teste-reteste foi complementada por meio da determinação do *standard error of measurement*, ou seja, erro padrão da medida (EPM) e do CV. O EPM foi calculado por meio da divisão do DP das diferenças entre o tempo do teste e reteste pela raiz quadrada de dois ( $\sqrt{2}$ ) (Hopkins, 2000; Hopkins et al., 2001; Batterham e Atkinson, 2005; Hopkins, 2009a). O CV foi determinado obtendo-se primeiramente EPM do logaritmo natural das variáveis ( $EPM_{ln}$ ), ou seja, por meio da divisão do DP das diferenças entre o logaritmo natural do tempo do teste e reteste pela  $\sqrt{2}$ . Evitamos assim problemas de heterocedasticidade (Hopkins, 2009a). Feito isso, o CV foi calculado por meio da fórmula  $CV = [\exp(EPM_{ln}) - 1]$ , em que  $\exp$  é a função exponencial natural (Hopkins, 2009a). Esses procedimentos para os cálculos do EPM e CV foram escolhidos por terem sido propostos por Hopkins (2009a) e constarem em planilha de acesso livre disponível na internet, com o cálculo dos respectivos intervalos de confiança (Hopkins, 2009b). Adotou-se o nível de significância de  $p < 0,05$ .

## Resultados

A tabela 1 apresenta os valores médios ± DP e a análise de Bland-Altman (1986) dos tempos das provas de 5 e 10 km (min:s) obtidos nos testes em pista de atletismo. Não houve diferença significativa entre os testes-retestes para ambas as provas ( $p > 0,05$ ). Em termos absolutos, verifica-se que a amplitude dos limites de concordância entre as provas de 10 km ( $\pm 1,96 \times 01:40$  min:s) foi maior do que a amplitude entre as provas de 5 km ( $\pm 1,96 \times 01:02$  min:s). Entretanto, em termos de proporção (limites de concordância/tempo

**Tabela 1** Valores médios  $\pm$  DP e análise de Bland-Altman dos tempos (min:s) das provas de 5 e 10 km ( $n = 31$ )

Prova	Média $\pm$ DP		Média $\pm$ DP das diferenças	Limites de concordância (95%)
	Teste	Reteste		
5 km	21:15 $\pm$ 02:51	21:08 $\pm$ 02:55	-00:07 $\pm$ 01:02	-02:11-01:58
10 km	44:44 $\pm$ 06:36	44:48 $\pm$ 07:05	00:03 $\pm$ 01:40	-03:16-03:22

Nota: os limites de concordância (95%) representam a média  $\pm$  1,96 DP das diferenças entre o teste-reteste.

**Tabela 2** Índices de reprodutibilidade teste-reteste das provas de 5 e 10 km representados pelo coeficiente de correlação de Pearson ( $r$ ), coeficiente de correlação intraclassa (CCI), erro padrão da medição (EPM) e coeficiente de variação (CV) para  $n = 31$ 

Prova	$r$	CCI (IC 95%)	EPM (min:s)	CV (IC 95%)
5 km	0,94	0,94 (0,87-0,97)	00:44	3,44% (2,74-4,63)
10 km	0,97	0,97 (0,94-0,99)	01:10	2,43% (1,93-3,26)

IC 95%, intervalo de confiança de 95%.

médio das provas), a amplitude dos limites de concordância entre as provas de 5 km ( $\pm$  9,6%) foi maior do que a amplitude entre as provas de 10 km ( $\pm$  7,3%).

A [tabela 2](#) apresenta os índices de reprodutibilidade teste-reteste do tempo das provas de 5 e 10 km. Observa-se que ambas as provas apresentaram uma alta reprodutibilidade representada pelo  $CCI > 0,90$ . Em relação ao EPM, o qual representa aproximadamente 68% do erro no valor da variável medida, foi menor na prova de 5 km em relação à prova de 10 km. Já o CV, que leva em consideração o tempo médio da prova, foi bem mais elevado na prova de 5 km do que na prova de 10 km.

A [figura 1](#) apresenta os gráficos de [Bland e Altman \(1986\)](#) para o tempo das provas de 5 km (esquerda) e 10 km (direita) em que o viés pode ser considerado zero, visto não ter havido diferença significativa ( $P > 0,05$ ) entre os testes-retestes para ambas as provas. Apesar da amplitude dos limites de concordância ser maior entre as provas de 10 km do que entre as provas de 5 km, o CV foi menor na prova de 10 km ([tabela 2](#)). Além disso, ainda que os dois pontos extremos nas provas de 10 km se apresentem como *outliers*, não foi observada anormalidade durante os testes, razão pela qual entendemos que esses dois pontos refletiram variações naturais dos participantes durante as provas. Assim, decidimos por não excluir esses dois pontos da análise.

A [tabela 3](#) apresenta o  $n_{\text{sign}}$  necessário para se detectarem os respectivos aumentos percentuais no desempenho ( $D$ ) com um nível de significância de 5% e um poder estatístico de 80%. Verifica-se, por exemplo, que para uma amostra de 10 participantes a prova de 10 km, ao contrário da prova de 5 km, provavelmente detectaria aumentos de 3% no desempenho. Já para aumentos inferiores a 1%, os quais seriam muito importantes para corredores de elite, ambas as provas se mostram inadequadas, a menos que um pesquisador disponha de pelo menos 48 corredores de elite dispostos a participar de um estudo.

## Discussão

O principal achado deste estudo foi que a reprodutibilidade do desempenho de corrida em provas de 5 e 10 km

em pista de atletismo foi semelhante às reprodutibilidades reportadas na literatura para testes em laboratório ([Doyle e Martinez, 1998](#); [Schabort et al., 1998b](#); [Rollo et al., 2008](#); [Russel et al., 2004](#)). Além disso, a prova de 10 km se mostrou mais reprodutível do que a prova de 5 km para corredores recreacionistas treinados do gênero masculino e necessitou de aproximadamente metade da amostra necessária para se detectarem aumentos significativos durante teste-reteste em provas de 5 km.

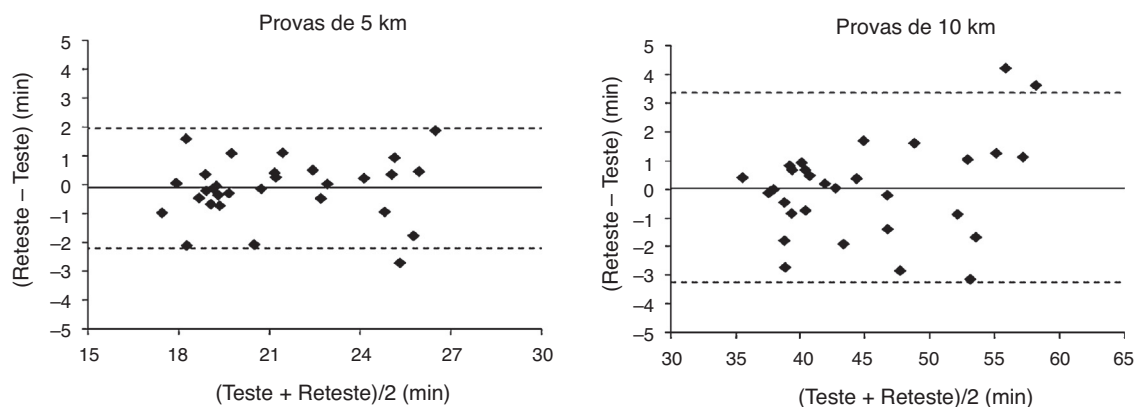
Para se examinarem os efeitos do treinamento ou de recursos ergogênicos sobre o desempenho é necessário que a variável usada para se determinar o desempenho seja altamente reprodutível ([Atkinson e Nevill, 1998](#)). Neste estudo, a variável de desempenho usada foi o tempo para se percorrerem 5 ou 10 km em pista de atletismo e os valores encontrados dessa reprodutibilidade estão de acordo com outros estudos que relataram um CV entre 1% e 4,4% ([Doyle e Martinez, 1998](#); [Schabort et al., 1998b](#); [Rollo et al., 2008](#); [Russel et al., 2004](#)). Há de ser enfatizar que esses estudos analisaram a reprodutibilidade do tempo para se percorrer certa distância ou a reprodutibilidade da maior distância percorrida em certo tempo. No entanto, ainda que sejam variáveis diferentes, tanto o tempo como a distância percorrida ou velocidade média representam CV idênticos durante provas de corrida ([Hopkins et al., 2001](#)) e, portanto, todas essas diferentes variáveis podem ser usadas de forma

**Tabela 3** Tamanho da amostra necessário para se verificarem aumentos significativos ( $n_{\text{sign}}$ ) no tempo das provas de 5 e 10 km para os respectivos aumentos percentuais no desempenho ( $D$ )

$D$	$N_{\text{sign}}$	
	Provas de 5 km	Provas de 10 km
1,0%	94	48
2,0%	25	14
3,0%	13	7
4,0%	8	5

Nota: considerando um teste unilateral com nível de significância de 5% e um poder estatístico de 80%.





**Figura 1** Gráficos de Bland-Altman entre os tempos dos testes-retestes para as provas de 5 km (esquerda) e 10 km (direita). A linha contínua representa o viés (média das diferenças entre os testes) e as linhas tracejadas representam os limites de concordância (95%) inferior e superior, calculados como o viés  $\pm 1,96 \times DP$  das diferenças entre os testes.

indistinta para se avaliar o desempenho. Não são necessários ajustes dos CV para as suas comparações. Ainda que apenas tenhamos informado o CV do tempo das provas neste estudo, calculamos o CV da velocidade média das provas e encontramos, conforme esperado, resultado idêntico ao do CV do tempo, ou seja, 3,44% (5 km) e 2,43% (10 km).

Para testes em laboratório, [Rollo et al. \(2008\)](#) reportaram um CV bastante baixo (1,4%) para a distância percorrida durante 60 minutos com o uso de uma esteira ergométrica em que a velocidade de corrida era controlada automaticamente, sem a necessidade de ajustes manuais. Já [Schabort et al. \(1998b\)](#) encontraram um CV de 2,7% para a distância percorrida durante 60 minutos, mas os participantes ajustavam a velocidade da esteira manualmente. O protocolo usado por [Russel et al. \(2004\)](#) foi altamente reprodutível (CV=1%). No entanto, demandou uma corrida preliminar de 90 minutos a 65% do  $VO_2\max$ , seguido pela reprodutibilidade do tempo para se percorrerem 10 km. Esse último protocolo demanda muita motivação e tem uma duração maior do que duas horas, o que o torna inadequado para ser aplicado em atletas que competem em eventos com duração aproximada de uma hora ([Doyle e Martinez, 1998](#); [Rollo et al., 2008](#)). Nosso estudo foi conduzido em pista de atletismo, com as provas se iniciando no mesmo horário sob condições ambientais semelhantes, e encontramos para a reprodutibilidade do tempo em provas de 10 km um resultado levemente inferior ao do estudo de [Schabort et al. \(1998b\)](#). Diante disso, ainda que as condições ambientais em pistas de atletismo não possam ser controladas tal como em laboratório, o que torna uma desvantagem a sua aplicação, esse fato é aparentemente compensado pela liberdade na qual os corredores podem alterar o ritmo de corrida livre e espontaneamente. Não tem a desvantagem dos ajustes manuais da velocidade que ocorrem em esteiras ergométricas tradicionais durante testes em laboratório ([Rollo et al., 2008](#)).

Outro ponto em relação ao CV relatado nesses estudos diz respeito ao seu método de determinação. Em geral, esses estudos determinaram o CV em amostras pequenas, em que os participantes fizeram três testes e assim determinaram a variabilidade intrassujeito com o auxílio da Anova para medidas repetidas. Neste estudo, usamos apenas dois testes e determinamos o CV de acordo com [Hopkins \(2009a\)](#), visto

que a Anova somente é apropriada para três ou mais testes ([Hopkins et al., 2001](#)). Ainda que existam outras formas para sua determinação, o procedimento usado neste estudo apenas alteraria o CV por um fator  $\geq 1,04$  nos casos raros, o qual não se enquadra neste estudo, em que o DP da amostra estudada é maior do que 35% da média dessa amostra ([Hopkins et al., 2001](#)). Ainda assim, para efeitos de comparação, podemos determinar o CV de outra maneira a partir dos dados das [tabelas 1 e 2](#) considerando a fórmula  $CV = [DP \times (1 - CCI)^{1/2}] / \text{média}$  que determina o CV por meio da divisão do erro típico da variável pela média dessa variável ([Atkinson e Nevill, 1998](#); [Hopkins et al., 2001](#)). Para a prova de 5 km, por exemplo, arredondando o DP para 3 minutos, a média para 21 minutos e o CCI para 0,94 teríamos de forma aproximada,  $CV = [3 \times (1 - 0,94)^{1/2}] / 21 = 3,5\%$ , bastante próximo do valor encontrado de 3,43%. Diante disso, com a devida cautela na aplicação dos métodos, esses não devem causar muito viés ao valor do CV.

Para a feitura de testes de reprodutibilidade, além do controle de variáveis como dieta, ausência de lesões, temperatura ambiente e horário, é importante que os atletas estejam familiarizados com o procedimento de avaliação ([Russel et al., 2004](#)) e encontrem-se motivados para os testes ([Jeukendrup et al., 1996](#)). [Doyle e Martinez \(1998\)](#) observaram que a reprodutibilidade do tempo para a distância média de 5,05 km melhorou drasticamente de um CV de 10,1% para um CV de 4,4% quando o primeiro teste foi considerado de aprendizado/familiarização e excluído das análises. Em outro estudo, [Laursen et al. \(2007\)](#) verificaram a reprodutibilidade das provas de 1,5 e 5 km em corredores bem treinados e, de acordo com seus resultados, a prova de 5 km foi mais reprodutível do que a prova de 1,5 km. Segundo os autores, isso pode estar relacionado à maior experiência e familiarização com a prova mais longa, tendo em vista que os corredores percorriam normalmente distâncias maiores do que 1,5 km em suas sessões de treinamento e competições. Neste estudo, os corredores recreacionistas estavam acostumados com as principais provas de rua da região, que são em geral iguais ou superiores à distância de 10 km. Em uma primeira análise, essa maior familiarização explicaria o menor CV da prova de 10 km comparada com a prova de 5 km e até poderia nos levar a concluir que o CV da prova de 5 km seria menor caso tivesse sido feita uma prova de familiarização.

Entretanto, como não houve uma diminuição significativa no tempo das provas de 5 km, não é possível afirmar que o CV encontrado está superestimado para essa população. Há de se ressaltar que as provas de 10 km foram feitas na semana subsequente às provas de 5 km, e não de forma aleatória. Essa falha metodológica deve ser considerada uma limitação deste estudo que pode ter influenciado no CV da prova de 5 km.

O conhecimento do CV de um teste é de suma importância quando se pretende estudar os efeitos do treinamento ou de recursos ergogênicos sobre o desempenho. Rollo et al. (2011), por exemplo, relataram um aumento de 2,2% na distância percorrida durante uma prova de corrida de uma hora após uma suplementação com carboidratos. Como base nos resultados desse estudo (tabela 3), 14 sujeitos seriam suficientes para se detectar esse aumento de 2,2% com o uso de um teste-reteste em provas de 10 km em pista de atletismo. Da mesma forma, seria necessário um número amostral superior a 48 sujeitos para se detectarem aumentos menores do que 1% no tempo entre as provas de 10 km em pista de atletismo, número esse que provavelmente inviabilizaria um estudo. Pequenas mudanças, da ordem de 1%, são bastante significativas para atletas de elite e levam um atleta a vencer ou perder uma prova, já que isso representa aproximadamente 15 segundos em uma prova de corrida de 10 km (Schabort et al., 1998b). Dessa forma, estudos que pretendam encontrar aumentos da ordem 1% devem usar testes com uma maior reprodutibilidade do que o apresentado neste estudo, como por exemplo, um teste de 60 minutos em esteira ergométrica com ajuste automático da velocidade tal como proposto por Rollo et al. (2008), em que as mudanças de velocidade da esteira são feitas sem a necessidade de ajustes manuais ou contato visual do corredor por meio de um sensor de posição, de acordo com os movimentos para frente ou para trás desse corredor (Minetti et al., 2003).

Há ainda que se ressaltar que o tamanho da amostra é dependente do CV e esse influenciado pela característica da população. Além do número de sujeitos, outros fatores devem ser levados em conta na composição das amostras, tal como o gênero, a idade dos sujeitos e o nível de competitividade (i.e, atleta x não atletas). Segundo Russell et al. (2004) e Hopkins e Hewson (2001), tanto os atletas mais velhos como os mais rápidos apresentam resultados mais reprodutíveis do que os seus congêneres mais novos e mais lentos. Hopkins et al. (2001) reportaram que o CV é aproximadamente 1,3 vez maior para não atletas do que atletas. Além disso, reportaram que o CV das mulheres não atletas é aproximadamente 1,4 vez maior do que o CV dos homens não atletas. Neste estudo, a idade avançada dos participantes provavelmente contribuiu para um menor CV, ao passo que o baixo nível de competitividade, uma vez que os participantes eram recreacionistas e competiam localmente, provavelmente contribuiu para um maior CV. Diante disso, é preciso cautela na aplicação dos resultados encontrados neste estudo em outras populações de diferentes idades, gênero, ou nível de competitividade.

Um último item que merece atenção diz respeito às variáveis usadas para se analisar a reprodutibilidade. Um consenso geral é que coeficientes de correlação avaliam a associação entre duas variáveis, informam o quanto a

posição relativa de cada sujeito varia junto entre o teste e o reteste, mas não detectam erros sistemáticos e são sensíveis à heterogeneidade, i.e, variabilidade interindividual (Atkinson e Nevill, 1998; Bland e Altman, 1986; Hopkins et al., 2001). Assim, é possível que o teste e reteste sejam fortemente correlacionados, mas não altamente reprodutíveis. Em relação ao EPM, CV e aos limites de concordância de Bland e Altman (1986), enquanto Hopkins et al. (2001) preferem os primeiros e mencionam que os limites de concordância de 95% são muito amplos e dificultam a tomada de decisão, Atkinson e Nevill (1998) preferem os limites de concordância de 95%, visto que o EPM e o CV abrangem apenas 68% da amostra. De fato, essas três variáveis podem ser convertidas entre si. Sabe-se que o CV pode ser derivado do EPM dividido pela média dos testes e sabe-se que os limites de concordância equivalem a  $\pm 2,77$  vezes o EPM (Atkinson e Nevill, 1998; Hopkins et al., 2001). Dessa forma, enquanto não houver um consenso entre os estudos, o melhor a ser feito é apresentar mais do que uma variável de reprodutibilidade para que esses diferentes estudos possam ser comparados.

## Conclusão

A reprodutibilidade encontrada para o desempenho em provas de corrida de 5 e 10 km em provas de atletismo se mostrou aceitável, semelhante às reportadas na literatura para testes em laboratório, ainda que não tenha sido suficiente para se detectarem aumentos no desempenho da ordem de 1% sem que necessite de um número muito elevado de sujeitos. Além disso, a prova de 10 km se mostrou mais reprodutível do que a prova de 5 km para corredores recreacionistas treinados do gênero masculino e necessitou de aproximadamente metade da amostra necessária para se detectarem aumentos significativos durante teste-reteste em provas de 5 km. Os achados deste estudo contribuem com a literatura por ser o primeiro a avaliar a reprodutibilidade do desempenho de *endurance* em pista de atletismo, ambiente mais próximo do onde realmente ocorrem as competições. Até então, os estudos que avaliaram essas variáveis foram feitos em ambiente laboratorial. Diante disso, recomendamos o uso de provas de 10 km, mais reprodutíveis do que provas de 5 km, para a avaliação de programas de treinamento ou efeitos de recursos ergogênicos sobre o desempenho de corredores recreacionistas treinados do gênero masculino. Futuros estudos poderão avaliar a reprodutibilidade em pista de atletismo tanto em outras populações de diferentes idades, gênero ou nível de competitividade como para outras distâncias ou diferentes durações de provas (e.g., 30 min; 1 h).

## Conflitos de interesse

Os autores declaram não haver conflitos de interesse.

## Referências

- Atkinson G, Nevill AM. *Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med* 1998;26(4):217–38.

- Batterham A, Atkinson G. How big does my sample need to be? A primer on the murky world of sample size estimation. *Phys Ther Sport* 2005;6(3):153-63.
- Beelen M, Berghuis J, Bonaparte B, Ballak SB, Jeukendrup AE, van Loon LJ. Carbohydrate mouth rinsing in the fed state: lack of enhancement of time-trial performance. *Int J Sport Nutr Exerc Metab* 2009;19(4):400-9.
- Bishop D. Reliability of a 1-h endurance performance test in trained female cyclists. *Med Sci Sports Exerc* 1997;29(4):554-9.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986 Feb 8;1(8476):307-10.
- Carter JM, Jeukendrup AE, Jones DA. The effect of carbohydrate mouth rinse on 1-h cycle time trial performance. *Med Sci Sports Exerc* 2004;36(12):2107-11.
- Chen Y, Wong SHS, Xu X, Hao X, Wong CK, Lam CW. Effect of CHO loading patterns on running performance. *Int J Sports Med* 2008;29(7):598-606.
- Doyle JA, Martinez AL. Reliability of a protocol for testing endurance performance in runners and cyclists. *Res Q Exercise Sport* 1998;69(3):304-7.
- Hickey MS, Costill DL, McConell GK, Widrick JJ, Tanaka H. Day to day variation in time trial cycling performance. *Int J Sports Med* 1992;13(6):467-70.
- Hopkins WG, Hawley JA, Burke LM. Design and analysis of research on sport performance enhancement. *Med Sci Sports Exerc* 1999;31(3):472-85.
- Hopkins WG. Measures of reliability in sports medicine and science. *Sports Medicine* 2000;30(1):1-15.
- Hopkins WG, Hewson DJ. Variability of competitive performance of distance runners. *Med Sci Sports Exerc* 2001;33(9):1588-92.
- Hopkins WG, Schabert EJ, Hawley JA. Reliability of power in physical performance tests. *Sports Med* 2001;31(3):211-34.
- Hopkins WG. A new view of statistics. mai. 2009a. Disponível em: <<http://www.sportsci.org/resource/stats/relycalc.html>>. Acesso em: 27 fev 2012.
- Hopkins W. Analysis of Reliability with a spreadsheet. mai. 2009b. Disponível em: <<http://www.sportsci.org/resource/stats/relycalc.html#excel>>. Acesso em: 27 fev 2012.
- Jensen K, Johansen L. Reproducibility and validity of physiological parameters measured in cyclists riding on racing bikes placed on a stationary magnetic brake. *Scand J Med Sci Sports* 1998;8(1):1-6.
- Jeukendrup A, Saris WH, Brouns F, Kester AD. A new validated endurance performance test. *Med Sci Sports Exerc* 1996;28(2):266-70.
- Laursen PB, Francis GT, Abbiss CR, Newton MJ, Nosaka K. Reliability of time-to-exhaustion versus time-trial running tests in runners. *Med Sci Sports Exerc* 2007;39(8):1374-9.
- Lindsay FH, Hawley JA, Myburgh KH, Schomer HH, Noakes TD, Dennis SC. Improved athletic performance in highly trained cyclists after interval training. *Med Sci Sports Exerc* 1996;28(11):1427-34.
- Madsen K, MacLean DA, Kiens B, Christensen D. Effects of glucose, glucose plus branched-chain amino acids, or placebo on bike performance over 100 km. *J Appl Physiol* 1996;81(6):2644-50.
- Minetti AE, Boldrini L, Brusamolín L, Zamparo P, McKee T. A feedback-controlled treadmill (treadmill-on-demand) and the spontaneous speed of walking and running in humans. *J Appl Physiol* 2003;95(2):838-43.
- Palmer GS, Dennis SC, Noakes TD, Hawley JA. Assessment of the reproducibility of performance testing on an air-braked cycle ergometer. *Int J Sports Med* 1996;17(4):293-8.
- Rollo I, Williams C, Nevill A. Repeatability of scores on a novel test of endurance running performance. *J Sports Sci* 2008;26(13):1379-86.
- Rollo I, Williams C. Influence of ingesting a carbohydrate-electrolyte solution before and during a 1-hour run in fed endurance-trained runners. *J Sports Sci* 2010;28(6):593-601.
- Rollo I, Williams C, Nevill M. Influence of ingesting versus mouth rinsing a carbohydrate solution during a 1-h run. *Med Sci Sports Exerc* 2011;43(3):468-75.
- Russell RD, Redmann SM, Ravussin E, Hunter GR, Larson-Meyer DE. Reproducibility of endurance performance on a treadmill using a preloaded time trial. *Med Sci Sports Exerc* 2004;36(4):717-24.
- Schabert EJ, Hawley JA, Hopkins WG, Mujika I, Noakes TD. A new reliable laboratory test of endurance performance for road cyclists. *Med Sci Sports Exerc* 1998a;30(12):1744-50.
- Schabert EJ, Hopkins WG, Hawley JA. Reproducibility of self-paced treadmill performance of trained endurance runners. *Int J Sports Med* 1998b;19(1):48-51.
- Smith MF, Davison RC, Balmer J, Bird SR. Reliability of mean power recorded during indoor and outdoor self-paced 40 km cycling time-trials. *Int J Sports Med* 2001;22(4):270-4.
- Whitham M, Mckinney J. Effect of a carbohydrate mouthwash on running time-trial performance. *J Sports Sci* 2007;25(12):1385-92.