

Reprodutibilidade no Teste de Uma Repetição Máxima no Exercício de Puxada Pela Frente Para Homens

ARTIGO ORIGINAL



Reproducibility in the Maximal Repetition Test in the Lat Pulldown Exercise For Men

Marcos André Pereira de Barros¹
Sandro Sperandei²
Paulo César Soter Silveira Júnior¹
Carlos Gomes de Oliveira³

1. Instituto de Ciências da Atividade Física – Rio de Janeiro / Brasil.

2. Programa de Engenharia Biomédica – COPPE/UFRJ – Rio de Janeiro / Brasil.

3. Programa de Pós Graduação em Ciências da Atividade Física - Universidade Salgado de Oliveira – Niterói / Brasil.

Endereço para correspondência:

Sandro Sperandei
Rua Tejo, 46 – Vila Valqueire
21321-300 – Rio de Janeiro, RJ
E-mail: ssperandei@hotmail.com

Submetido em 30/12/2005
Versão final recebida em 07/11/2006
Aceito em 21/11/2006

RESUMO

Objetivo: Verificar a reprodutibilidade teste-reteste do teste de uma repetição máxima (1RM) no exercício de puxada pela frente. **Métodos:** 26 homens adultos saudáveis e ativos (25 ± 3 anos, $76,8 \pm 9,86$ kg, $177 \pm 5,3$ cm, $12,5 \pm 4,25\%$ de gordura) assinaram consentimento antes de participar deste estudo. A coleta de dados foi feita em dois dias, separados por 48-72 horas. No primeiro dia, foram coletados dados antropométricos e o teste de 1RM progressivo foi realizado. Em cada tentativa, cargas progressivas de 5kg eram acrescentadas até que ocorresse uma falha, quando então a carga era reduzida em três quilos e uma nova tentativa era realizada. Foram respeitados cinco minutos de intervalo entre as tentativas. No segundo dia, o teste de 1RM era repetido seguindo o mesmo protocolo. As cargas máximas dos dois dias foram comparadas utilizando o teste *t* pareado para o erro sistemático e a técnica de limites de concordância, para 95% do erro aleatório. **Resultados:** O erro médio apresentado pela amostra foi de $2,9 \pm 2,48$ kg, sendo observada diferença significativa entre os dois dias de teste ($p < 0,00001$). O limite de concordância para 95% do erro ficou entre -1,9 e 7,8kg. **Conclusão:** Este estudo mostrou tendência a aumento da carga máxima atingida no segundo dia de teste, o que aponta para a necessidade de se considerar tal diferença ao planejar e analisar estudos de força onde o teste de 1RM for utilizado.

Palavras-chave: reprodutibilidade, treinamento de força, limites de concordância.

ABSTRACT

Objective: To verify the test-retest 1RM reproducibility in the lat pulldown exercise. **Methods:** Twenty-six healthy and physically active men (25 ± 3 years, 76.8 ± 9.86 kg, 177 ± 5.3 cm, 12.5 ± 4.25 % body fat) signed an informed consent before participate in this study. The data was collected in two visits, 48-72 hours apart. During the first visit, anthropometric data was measured and progressive 1RM test was performed. For each attempt progressive loads of 5 kg were added until the fail, when the load was reduced in 3 kg and a last attempt was performed. A five-minute rest was allowed before each attempt. During the second visit, the 1RM test was repeated following the same protocol. The Maximum loads of each day were compared by paired t-test to systematic error and limits of agreement technique to random error was employed. **Results:** The mean error found was 2.9 ± 2.48 kg, being significantly different between days ($p < 0.00001$). The error 95% limit of agreement was between -1.9 and 7.8 kg. **Conclusion:** This study showed a tendency to an increase in the second-day load, which when planning and analyzing strength training studies using 1RM test.

Keywords: reproducibility, strength training, limits of agreement.

INTRODUÇÃO

O treinamento de força tem recebido especial atenção nos últimos anos. O *American College of Sports Medicine*, umas das principais instituições no mundo relacionadas ao exercício físico atualmente, destaca a importância do treinamento de força ao incluí-lo em seu posicionamento oficial sobre atividade física e saúde⁽¹⁾ e, posteriormente, publicar um posicionamento apenas relacionado ao tema⁽²⁾. Em 1997, Fleck e Kraemer⁽³⁾ afirmaram que haviam sido publicadas mais pesquisas sobre o assunto entre 1987 e 1997 do que nos 50 anos anteriores. Desde então, o número de pesquisas na área vem aumentando ainda mais.

O teste de Uma Repetição Máxima (1RM) é o principal método utilizado para aferir a força de um indivíduo, consistindo, basicamente, em encontrar a carga máxima que um indivíduo é capaz de utilizar ao completar uma única repetição de determinado exercício com execução correta.

O uso de diferentes procedimentos durante a realização do teste pode influenciar de maneira direta o seu resultado final. Por exemplo, incrementos de carga muito grandes podem levar a subestimação da carga máxima, enquanto incrementos muito reduzidos ocasionam grande número de tentativas. Outras variáveis que precisam ser controladas são o intervalo de recuperação entre as tentativas e a amplitude de movimento. Diferentes protocolos de teste para

1RM podem ser encontrados na literatura específica, mas nenhum que tenha padronizado os procedimentos de maneira adequada, determinando valores exatos para incrementos de carga, intervalos de recuperação adequados e formas de controle da amplitude de movimento. Por este fato, neste trabalho, utilizamos um protocolo desenvolvido em nosso grupo de pesquisa, descrito na seção de Métodos.

O teste de 1RM já foi utilizado como forma de determinar a intensidade do treinamento, no qual a carga era expressa em percentuais da carga para 1RM. No entanto, essa utilização caiu em desuso devido à dificuldade de se realizar o teste periodicamente e à grande variabilidade no número de repetições possíveis em diferentes exercícios com um mesmo percentual da carga máxima⁽³⁾. Atualmente, o teste é utilizado antes e após um treinamento, como forma de monitorar as variações possivelmente relacionadas ao protocolo de treino aplicado. Outra forma de utilização é para estabelecer uma padronização na intensidade da carga a ser utilizada durante uma investigação científica. Nesse caso, o teste é realizado e um dado percentual do resultado é utilizado na realização do exercício.

Considerando os objetivos acima descritos, a confiabilidade do teste de 1RM precisa ser muito boa, de forma que a variação encontrada entre duas medidas seja resultado de alguma forma de treinamento aplicado ou fato similar, e não um erro da medida, e também para que a intensidade utilizada seja realmente aquela esperada. Por essa razão, a reprodutibilidade do teste de 1RM tem sido objeto de estudo de alguns pesquisadores, mas não parece ter sido suficientemente investigada. Em uma revisão sobre o assunto, Pereira e Gomes⁽⁴⁾ encontraram apenas quatro estudos publicados abordando o tema. Todos os estudos apresentados relataram confiabilidade alta, utilizando a correlação de Pearson (*r*) ou o coeficiente de correlação intraclasse (ICC) como método para aferir essa confiabilidade. Tais métodos apresentam importantes limitações, como já amplamente discutido na literatura⁽⁵⁻⁷⁾. Além das limitações matemáticas, os resultados desses estudos se mostram de pouca validade prática, devido à impossibilidade de avaliar o tamanho absoluto do erro da medida a partir de um coeficiente⁽⁵⁾.

Os métodos mais adequados para a estimação da reprodutibilidade de um teste parecem ser o erro típico (*typical error*)⁽⁶⁾ e o limite de concordância (*limits of agreement*)^(5,7). No entanto, apenas dois trabalhos foram encontrados utilizando tais métodos para medir a confiabilidade do teste de 1RM⁽⁸⁻⁹⁾, mas nenhum avaliou exercícios que fossem pelo menos similares ao da puxada pela frente.

Dessa forma, o objetivo deste estudo foi verificar a reprodutibilidade do teste de uma repetição máxima (1RM), na situação teste-reteste, no exercício de puxada pela frente em indivíduos do sexo masculino.

MÉTODOS

Foram voluntários 26 indivíduos adultos jovens (25 ± 3 anos, $76,8 \pm 9,86$ kg, $177 \pm 5,3$ cm, $12,5 \pm 4,25\%$ de gordura), do sexo masculino, aparentemente saudáveis e sem histórico de dor prolongada ou lesão nas articulações do ombro e cotovelo nos seis meses que antecederam a coleta dos dados.

Todos os voluntários possuíam experiência prévia de, no mínimo, seis meses no treinamento de força utilizando o exercício selecionado ou similar e assinaram um termo de consentimento antes de participar do estudo, de acordo com a Resolução 196/96 do Conselho Nacional de Saúde.

A coleta dos dados foi realizada em duas visitas, separadas por um mínimo de 48 e um máximo de 72 horas. Os voluntários foram instruídos a abster-se do treinamento de força para membros superiores e

tronco a partir de 48 horas antes do primeiro dia de coleta, devendo permanecer sem treinamento até o final da segunda visita.

Na primeira visita eram coletados a massa corporal, estatura, dobras cutâneas (Jackson e Pollock⁽¹⁰⁾) e distância biacromial, tendo esta última medida o objetivo de padronizar a posição de pegada na barra do exercício. A distância entre as mãos foi estabelecida em 1,6 vez a distância biacromial, e era marcada e registrada (figura 1). Após esse procedimento, o indivíduo era orientado a realizar uma série de 12 a 15 repetições, como forma de aquecimento, utilizando uma carga próxima a 50% daquela normalmente empregada durante o treinamento. Tanto o aquecimento como as subseqüentes tentativas do teste foram realizadas em um puxador alto padrão.

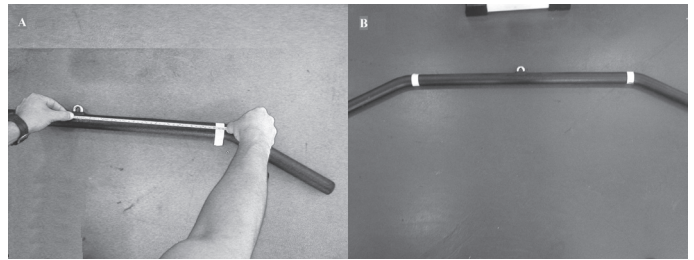


Figura 1. A - Marcação da posição de pegada na barra. B - Posição de pegada marcada.

Um intervalo entre as repetições de cinco minutos entre as tentativas era observado, durante o qual era realizada uma única repetição do exercício, sem sobrecarga, com o objetivo de estabelecer a amplitude do movimento. A barra era puxada até que ultrapassasse o plano horizontal do queixo do voluntário. Uma fita adesiva era presa ao aparelho, no ponto mais alto alcançado pelas placas de sobrecarga (figura 2). Essa medida era anotada e registrada para utilização também na segunda visita.

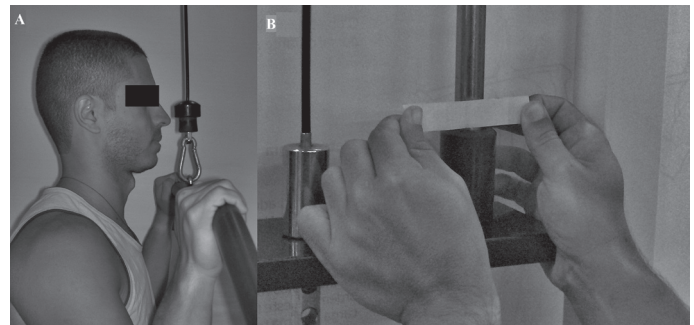


Figura 2. A - Barra na posição de amplitude máxima. B - Marcação da amplitude de movimento no aparelho com o indivíduo na posição A.

Após os cinco minutos de intervalo, o indivíduo realizava a sua primeira tentativa no teste de 1RM. Uma carga aproximadamente 30% superior àquela relatada pelo voluntário para 10RM (carga de treino) era utilizada nessa primeira tentativa. Com o indivíduo sentado, a barra era trazida ao seu alcance por um ajudante e mantida assim até que o próprio voluntário desse a ordem de início do teste. A partir desse momento, a barra era solta pelo ajudante e deveria ser tracionada pelo voluntário até que a primeira placa de sobrecarga alcançasse a fita presa ao aparelho. Nesse ponto, um aviso era dado pelo avaliador e o indivíduo deveria voltar a barra à posição original.

Em caso de sucesso, cinco quilos eram adicionados à carga e uma nova tentativa era realizada após o intervalo. As tentativas se sucediam até que uma falha na execução ocorresse. Nesse ponto, três quilos

eram retirados da carga e uma última tentativa era feita. A maior carga utilizada para uma tentativa realizada com sucesso era considerada a carga para 1RM.

Entre cada tentativa foi respeitado um intervalo de cinco minutos. Um máximo de seis tentativas era permitido em cada visita. Apenas um voluntário não alcançou a carga de 1RM e este teve que retornar outro dia para uma nova sessão. Na segunda visita, os procedimentos relativos ao teste de 1RM eram repetidos. Todos os testes foram aplicados pelo mesmo avaliador.

Os dados foram analisados com relação a sua diferença entre os dias (erro). Para a determinação da existência de erro sistemático entre os dias foi aplicado o teste *t* para amostras pareadas, utilizando-se $\alpha = 0,05$. Na identificação do erro aleatório entre os dias, foi aplicada a técnica de limites de concordância⁽⁷⁾, para o intervalo de 95% do erro. O coeficiente de Pearson e o ICC foram calculados para comparação com outros estudos publicados na literatura.

RESULTADOS

Os resultados médios dos dois dias de medida (tabela 1) mostraram aumento significativo na carga obtida no segundo dia ($p < 0,00001$).

Tabela 1. Resultados do teste de 1RM no primeiro e segundo dia, em quilos (média \pm DP).

	Dia 1	Dia 2	Erro	CI (95%)*
Carga 1RM	97,7(13,87)	100,6(14,46)	2,9(2,48)	1,92 – 3,93

*CI – Intervalo de confiança do erro médio.

Com relação aos limites de concordância, foi observado que, em 95% dos casos, o erro aleatório situou-se entre os valores de -1,9 e 7,8kg (figura 3), com $r = 0,985$ e ICC = 0,984.

DISCUSSÃO

O conhecimento do erro apresentado por um teste ou instrumento é de fundamental importância no momento de interpretar os resultados obtidos por um procedimento. A presença de um erro grande de medida no teste de 1RM pode, por exemplo, induzir a conclusões errôneas sobre a eficiência de um tipo de treinamento.

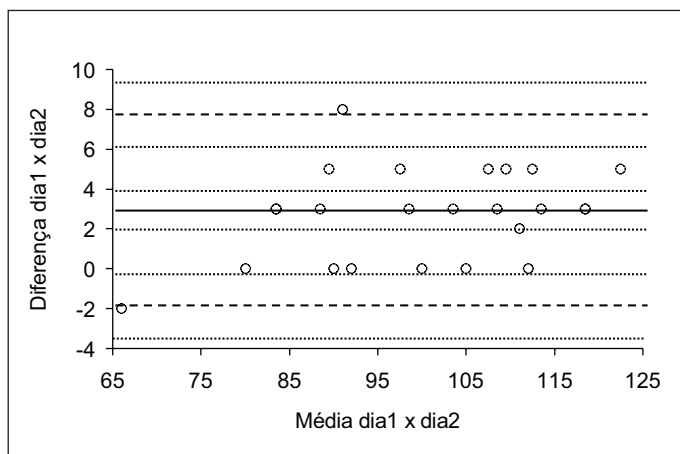


Figura 3. Bland-Altman plot. Gráfico da diferença (erro) entre os dias pela média entre os dias. A linha cheia indica a média do erro. As linhas tracejadas indicam os limites de concordância para 95% do erro. Não foi observada correlação significativa entre o erro e o tamanho da medida ($r=0,238$). Linha cheia – erro médio. Linhas tracejadas – limites de concordância. Linhas pontilhadas – intervalos de confiança dos parâmetros.

É também importante que se faça a distinção entre dois tipos de erro que podem estar presentes: erro sistemático e erro aleatório. O primeiro é caracterizado como um desvio a partir do valor real da medida que se apresenta sistematicamente na mesma direção e com magnitude mais ou menos regular. O erro aleatório, por outro lado, alterna em direção (erro para mais ou para menos) e varia em magnitude, sendo menos previsível.

Antes de serem feitas considerações sobre os resultados encontrados, torna-se importante destacar que o protocolo de teste apresentado aqui foi desenvolvido pelo próprio grupo de pesquisa, a partir da análise de diferentes protocolos descritos na literatura. Isso foi necessário devido à falta de padronização observada em testes de 1RM em termos de carga e incrementos. Dias *et al.*⁽¹¹⁾ e Ploutz-Snyder e Giamis⁽¹²⁾, por exemplo, não descrevem o incremento de carga utilizado, enquanto Braith *et al.*⁽⁸⁾ e Hoeger *et al.*⁽¹³⁾ utilizaram entre 2,3 e 9kg e Philips *et al.*⁽⁹⁾ variaram o incremento de acordo com a sensação de esforço do avaliado. Dessa forma, a utilização de incrementos diferentes por diferentes avaliadores pode interferir no erro de medida do teste. Uma das principais vantagens do protocolo utilizado aqui está na maior objetividade, facilitando a sua reprodução por outros avaliadores, quando, por exemplo, o incremento de carga utilizado é sempre de cinco quilos, equivalente a uma placa do aparelho utilizado. Outras vantagens do protocolo utilizado neste trabalho estão na utilização de uma tentativa final com redução da carga, a determinação precisa da posição de pegada na barra e o controle rígido da amplitude de movimento. No entanto, é fundamental destacar que os resultados obtidos de reprodutibilidade se aplicam unicamente aos testes realizados com o protocolo aqui descrito.

Os resultados do presente estudo demonstraram a presença dos dois tipos de erro, sistemático e aleatório, indicando tendência de que a segunda medida seja, em média, 2,9kg maior do que a primeira, e a variação, para 95% dos casos, fique entre -1,9kg e 7,8kg. Isso leva a crer que a utilização de incrementos menores entre as tentativas poderia levar a resultado um pouco mais preciso. Entretanto, esse procedimento demandaria número muito superior de tentativas, o que poderia causar interferência devido à fadiga. Outra possível limitação do presente estudo pode estar relacionada exatamente à fadiga. O *American College of Sports Medicine*⁽¹⁴⁾ recomenda um intervalo de três a cinco minutos entre cada tentativa. Matuszak⁽¹⁵⁾, por exemplo, não demonstrou diferença entre um, três e cinco minutos de intervalo, mas apenas para duas tentativas. O efeito da fadiga em cinco a seis tentativas consecutivas permanece por ser estabelecido. Acreditamos que o intervalo de cinco minutos é suficiente para a recuperação entre as tentativas, sendo o maior observado dentre os estudos analisados.

Dos estudos já publicados e analisados, apenas três^(8-9,16) relatam a presença de erro sistemático. No entanto, quando avaliados dentro das características do presente trabalho, outro autor⁽¹¹⁾ apresenta erro sistemático e mais um⁽¹²⁾ possivelmente também, ainda que não relatado de forma direta.

O trabalho de Dias *et al.*⁽¹¹⁾ demonstrou erro sistemático entre a primeira e a segunda tentativa para o teste realizado em três diferentes exercícios (1,1kg, 2,6kg e 1,5kg, para supino, agachamento e rosca bíceps, respectivamente), mas sua análise foi restrita ao “ponto de estabilização”, onde não houve diferença significativa entre as tentativas. Uma diferença importante entre a metodologia utilizada por esses autores e a aqui apresentada é que o teste era realizado de forma que a carga era aumentada apenas se o indivíduo fosse capaz

de realizar duas repetições consecutivas. Dessa forma, ainda que a segunda repetição fosse executada em 80% de sua amplitude, o que indicaria a possibilidade de aumento de carga, esta permaneceria a mesma.

De maneira semelhante, Ploutz-Snyder e Giamis⁽¹²⁾ relatam a necessidade de diversas tentativas (entre duas e 10) para alcançar a estabilização da carga no teste de 1RM. Os autores não descrevem as diferenças entre cada par de sessões consecutivas, mas relatam aumentos médios entre 12% e 22% nas cargas para última sessão em relação à primeira. Uma possível justificativa para a inexistência de erro sistemático, entre a penúltima e a última tentativa, é o critério que foi adotado para “estabilização”, qual seja, a diferença não superior a um quilo entre a carga para duas sessões consecutivas. No entanto, a necessidade de realização de um grande número de sessões de teste (seis sessões, em média) torna a aplicação dessa metodologia muito limitada à realidade prática da pesquisa, até pelo efeito de treinamento ocasionado, demonstrado pela já citada variação na carga entre a primeira e a última sessão.

Braith *et al.*⁽⁸⁾ reportaram erro sistemático de 3,9kg e 3,1kg, pré e pós um treinamento de força, representando 4,4% e 2,8% da carga para 1RM na extensão do joelho. Esses resultados são semelhantes aos encontrados no presente estudo, ainda que o intervalo de descanso entre as tentativas tenha sido inferior ao nosso.

Erro sistemático foi relatado também por Philips *et al.*⁽⁹⁾, entre a primeira e a segunda sessão de teste de 1RM, no supino realizado por homens. Seus resultados mostraram uma variação média de 5%, ou 4,9kg. No *leg press* não foi detectada diferença significativa, tanto para homens quanto para mulheres, que também não a apresentaram para o supino. Entre a segunda e terceira sessão foi observada diferença para mulheres no *leg press*, mas não no supino, nem para homens em nenhum dos dois exercícios.

Em um grupo de 42 indivíduos idosos, Rikli *et al.*⁽¹⁶⁾ encontraram erro sistemático entre o primeiro e o segundo dia de teste, mas não entre o segundo e o terceiro. Os autores descrevem ICCs variando entre 0,97 e 0,98 entre o segundo e o terceiro dia, mas não descrevem os valores encontrados para os erros sistemáticos. Entretanto, analisando as diferenças pré e pós-treinamento de 10 semanas, os autores relatam que o erro de medida equivale a percentual entre 31, 4% e 55,9% da diferença encontrada. Essa comparação mostra a importância da reprodutibilidade do teste e a relativa ineficiência do ICC em aferir essa variável.

O coeficiente de Pearson (*r*) e o coeficiente de correlação intraclass (ICC) vêm sendo utilizados de maneira geral como medida de reprodutibilidade do teste de 1RM. Ambos os métodos não apresentam sensibilidade ao erro sistemático, de tal forma que, mesmo que a carga na segunda sessão seja o dobro da primeira, tanto o ICC quanto o *r* serão iguais a 1,0, que indica perfeita reprodutibilidade.

Ao avaliar a reprodutibilidade de um teste, é comum encontrar a atenção voltada principalmente para o erro sistemático, que é identificado pela presença de diferença estatística significativa entre os dois momentos de aplicação do teste. A ausência de diferença significativa (calculada por um teste *t*, por exemplo) é utilizada como indicativo de boa confiabilidade para o teste. Atkinson e Nevill⁽⁵⁾ demonstraram a armadilha nesse tipo de interpretação com um exemplo simples, reproduzido na figura 4. Pode-se observar, nesse exemplo, que não há diferença significativa entre as duas séries de medidas (média da diferença igual a zero). No entanto, um teste que apresentasse tal resultado não poderia ser considerado “confiável”, dada a grande variação entre as duas medidas de cada “indivíduo”.

n	Medida 1	Medida 2	Diferença	$t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}}$	Onde: \bar{d} - Média da diferença; s_d - Desvio-padrão da diferença; n - Número de indivíduos; $t=0 \rightarrow p=1$ Sem diferença significativa!
1	1	10	9		
2	10	1	-9		
3	2	20	18		
4	20	2	-18		
5	3	30	27		
6	30	3	-27		
7	4	40	36		
8	40	4	-36		
Média	13,75	13,75	0,0		

Figura 4. Comparação de duas medidas utilizando o teste *t*. A inexistência de diferença significativa entre as medidas indica ausência de erro sistemático, mas não de erro aleatório (Adaptado de Atkinson e Nevill⁽⁵⁾).

O erro sistemático, apesar de comprometer em parte a confiabilidade do teste, é o mais fácil de ser solucionado, pois é um erro esperado e pode mesmo ser subtraído (ou adicionado) do resultado do teste subsequente. Por outro lado, o erro aleatório não demonstra um padrão definido e interfere diretamente na sensibilidade dos testes estatísticos para detectar diferenças entre dois tipos de intervenção.

Na avaliação do erro aleatório, o ICC apresenta pequena vantagem em relação ao *r*, como demonstrado por Ottenbacher e Tomcheck⁽¹⁷⁾, mas ainda insuficiente para a adoção do método, como será visto a seguir. Os valores de ICC e/ou correlação de Pearson descritos pelos trabalhos analisados e os encontrados no presente estudo podem ser observados na tabela 2. Todos os valores descritos para um ou outro coeficiente são considerados altos, indicando alta reprodutibilidade. Entretanto, mesmo para o erro aleatório, a capacidade desses dois métodos é muito limitada. Para uma revisão ampla das limitações, ver Atkinson e Nevill⁽⁵⁾, Hopkins⁽⁶⁾ e Altman e Bland⁽⁷⁾.

Além do presente estudo, dois outros utilizaram métodos reconhecidamente eficientes na avaliação do erro aleatório. Dias *et al.*⁽¹¹⁾ utilizaram o limite de concordância (LOA), mesmo método descrito aqui, para as sessões de estabilização, apesar de não descrever explicitamente os limites do erro aleatório. A partir dos gráficos apresentados, esses limites podem ser aproximados e são apresentados, para comparação, na tabela 3^(13,18).

Tabela 2. Valores de correlação (*r*) e de ICC (R) encontrados na literatura e no presente estudo

Estudo	R	r
Braith <i>et al.</i> – pós-treino ⁽⁸⁾	-	0,99
Braith <i>et al.</i> – pré-treino ⁽⁸⁾	-	0,98
Hoeger <i>et al.</i> – homens ⁽¹³⁾	0,89 - 0,98	
Hoeger <i>et al.</i> – mulheres ⁽¹³⁾	0,79 - 0,98	
Materko <i>et al.</i> – homens ⁽¹⁸⁾	-	0,99
Materko <i>et al.</i> – mulheres ⁽¹⁸⁾	-	0,94
Ploutz-Snyder e Giamis ⁽¹²⁾	-	0,97
Rikli <i>et al.</i> ⁽¹⁶⁾	0,97 - 0,98	
Presente estudo	0,984	0,985

Tabela 3. Valores de erro médio e limites de concordância (LOA) descritos na literatura e encontrados no presente estudo.

Estudo	Erro médio	LOA (95%)	
		Limite superior	Limite inferior
Dias <i>et al.</i> ⁽¹¹⁾ - supino - 3x4	0,5	2,3	-1,3
Dias <i>et al.</i> ⁽¹¹⁾ - agachamento - 3x4	1,4	7,1	-4,3
Dias <i>et al.</i> ⁽¹¹⁾ - rosca bíceps - 2x3	0,5	2,6	-1,6
Phillips <i>et al.</i> ⁽⁹⁾ - supino - homens - 1x2	4,9*	19,4	-9,7
Phillips <i>et al.</i> ⁽⁹⁾ - supino - homens - 2x3	0,7	13,4	-12,0
Phillips <i>et al.</i> ⁽⁹⁾ - leg press - homens - 1x2	-0,4	17,6	-18,4
Phillips <i>et al.</i> ⁽⁹⁾ - leg press - homens - 2x3	2,0	11,2	-7,3
Phillips <i>et al.</i> ⁽⁹⁾ - supino - mulheres - 1x2	1,1	15,1	-12,9
Phillips <i>et al.</i> ⁽⁹⁾ - supino - mulheres - 2x3	2,3	21,8	-17,2
Phillips <i>et al.</i> ⁽⁹⁾ - leg press - mulheres - 1x2	-0,1	16,8	-17,0
Phillips <i>et al.</i> ⁽⁹⁾ - leg press - mulheres - 2x3	3,5*	18,6	-11,6
Presente estudo	2,9*	7,8	-1,9

* $p < 0,05$

O erro típico (TE) é uma técnica válida na avaliação do erro aleatório⁽⁶⁾ e foi utilizada por Phillips *et al.*⁽⁹⁾. Dada a relação direta entre o TE e o LOA⁽¹⁹⁾, os valores apresentados por esses autores foram transformados e são apresentados também na tabela 3, junto aos valores encontrados no presente trabalho. Pode-se observar grande variabilidade nos limites encontrados para o erro aleatório, o que pode ser explicado por diferenças nos exercícios utilizados e na metodologia empregada na realização do teste.

O critério para 1RM utilizado por Dias *et al.*⁽¹¹⁾, já citado, parece ser responsável por boa parte da diferença no erro aleatório em relação ao apresentado neste estudo, pois torna o método menos sensível a variações. A realização de sessões de familiarização pode ser também responsável por parte das diferenças apresentadas em relação aos nossos resultados. No entanto, Phillips *et al.*⁽⁹⁾ realizaram no mínimo três sessões de familiarização, chegando a mais de seis, e encontraram um erro aleatório muito superior ao obtido neste estudo. Também, o nível de treinamento dos indivíduos avaliados pode ter influenciado, uma vez que Phillips *et al.*⁽⁹⁾ utilizaram idosos destreinados, enquanto em Dias *et al.*⁽¹¹⁾ e no presente trabalho eram jovens com um mínimo de seis meses de experiência.

CONCLUSÃO

O presente estudo observou a presença de erro sistemático e aleatório entre a primeira e a segunda sessão do teste de 1RM para o exercício de puxada pela frente. Apesar disso, o erro sistemático mostrou-se pequeno em relação ao valor absoluto da carga utilizada (aproximadamente 3%). O erro aleatório precisa ser levado em conta no planejamento e avaliação de estudos que utilizem o teste de 1RM como forma de aferir a força. Quanto maior o erro aleatório, maior deverá ser o número de indivíduos avaliados para a detecção de uma mesma diferença pré/pós-treinamento. Tais resultados nos levam a recomendar a realização de estudos de avaliação da reprodutibilidade, ainda que em forma de estudo-piloto, sempre antes da utilização do teste de 1RM como método de avaliação da força. Esse procedimento destina-se não só a orientar a avaliação dos resultados obtidos, como também a possibilitar melhor planejamento da pesquisa em si.

Todos os autores declararam não haver qualquer potencial conflito de interesses referente a este artigo.

REFERÊNCIAS BIBLIOGRÁFICAS

- American College of Sports Medicine Position Stand. The Recommended Quantity and Quality of Exercise for Developing and Maintaining Cardiorespiratory and Muscular Fitness, and Flexibility in Healthy Adults. *Med Sci Sports Exerc* 1998;30:975-91.
- Kraemer WJ, Adams K, Cafarelli E, Dudley GA, Dooly C, Feigenbaum MS, et al. American College of Sports Medicine Position Stand. Progression Models in Resistance Training for Healthy Adults. *Med Sci Sports Exerc* 2002;34:364-80.
- Fleck SJ, Kraemer WJ. Designing resistance training programs. 2nd ed. Champaign: Human Kinetics, 1997.
- Pereira MI, Gomes PS. Testes de força e resistência muscular: confiabilidade e predição de uma repetição máxima - Revisão e novas evidências. *Rev Bras Med Esp* 2003;9:325-35.
- Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med* 1998;26:217-38.
- Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med* 2000;30:1-15.
- Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. *The Statistician* 1983;32:307-17.
- Braith RW, Graves JE, Leggett SH, Pollock ML. Effect of training on the relationship between maximal and submaximal strength. *Med Sci Sports Exerc* 1993;25:132-8.
- Phillips WT, Batterham AM, Valenzuela JE, Burkett LN. Reliability of maximal strength testing in older adults. *Arch Phys Med Rehabil* 2004;85:329-34.
- Jackson AS, Pollock ML, Gettman LR. Intertester reliability of selected skinfold and circumference measurements and percent fat estimates. *Res Q* 1978;49:546-51.
- Dias RM, Cyrino ES, Salvador EP, Caldeira LF, Nakamura FY, Bruna N, et al. Influência do processo de familiarização para avaliação da força muscular em testes de 1-RM. *Rev Bras Med Esporte* 2005;11:34-8.
- Ploutz-Snyder LL, Giamis EL. Orientation and familiarization to 1RM strength testing in old and young women. *J Strength Cond Res* 2001;15:519-23.
- Hoeger WW, Hopkins DR, Barette SL, Hale DF. Relationship between repetitions and selected percentages of one repetition maximum: a comparison between untrained and trained males and females. *J Appl Sport Sci Res* 1990;4:47-54.
- American College of Sports Medicine, Whaley MH, Brubaker PH, Otto RM, Armstrong LE. ACSM's Guidelines for Exercise Testing and Prescription. 7th ed. Philadelphia, Pa: Lippincott Williams & Wilkins, 2006.
- Matuszak ME, Fry AC, Weiss LW, Ireland TR, McKnight MM. Effect of rest interval length on repeated 1 repetition maximum back squats. *J Strength Cond Res* 2003;17:634-7.
- Rikli RE, Jones CJ, Beam WC, Duncan SJ, Lamar B. Testing versus training effects on 1RM strength assessment in older adults [Resumo]. *Med Sci Sports Exerc* 1996;28:S153.
- Ottenbacher KJ, Tomchek SD. Measurement variation in method comparison studies: an empirical examination. *Arch Phys Med Rehabil* 1994;75:505-12.
- Materko W, Neves CE, Santos EL. Predição de uma repetição máxima (1RM) baseada nas características antropométricas [Resumo]. *Anais do XXVII Simpósio Internacional de Ciências do Esporte*, 2004;190.
- Atkinson G, Nevill A. Typical error versus limits of agreement. *Sports Med* 2000;30:375-81.