

## Reflexões sobre o viés de publicação: um guia para praticantes de estatística para a análise de dados e uso inapropriado do coeficiente de correlação em ciências da saúde

### Reflexions on publication bias: a guide for statisticians practitioners to data analysis and critical use of correlation coefficient in health sciences

Nascimento DC, Almeida JA, Carvalho THD, Prestes J. Reflexões sobre o viés de publicação: Um guia para praticantes de estatística para análise de dados e uso inapropriado do coeficiente de correlação em ciências da saúde. *R. bras. Ci. e Mov* 2020;28(4):194-201.

**RESUMO:** A utilização de procedimentos estatísticos é de fundamental importância para a interpretação apropriada de um conjunto de dados. Desta forma, a baixa aderência do teste aos dados selecionados pode levar a conclusões inadequadas. Portanto, a escolha do teste paramétrico e não paramétrico para dados pareados deve levar em conta a normalidade dos dados. Com isso, aplicar o coeficiente de correlação de *Pearson* (teste paramétrico) em dados não paramétricos aumenta as chances de associações espúrias (por acaso ou erro sistemático), as quais resultam em erro do Tipo I. Entendendo que as vezes o pensamento do jovem pesquisador e também de editores de periódicos científicos serão guiados por resultados positivos. É comum a possibilidade de editores selecionarem artigos para publicação tendo como base o valor de  $p < 0,05$ . Contudo, também seria importante selecionar os artigos levando em consideração os cumprimentos dos pressupostos para a utilização de testes paramétricos e não-paramétricos. Com isso, objetivo do presente estudo foi abordar os dois testes de coeficiente de correlação de *Pearson* e *Spearman* e sugerir recomendações para praticantes de estatística na área de Ciências da Saúde para a utilização segura e adequada dos dados antes da publicação

**Palavras-chave:** Correlação; Normalidade; Viés; Erro do tipo I.

**Abstract:** The use of statistical procedures is of fundamental importance for the proper interpretation of data analysis. In this way, the low adherence of the test to the selected data can lead to inadequate conclusions. Therefore, the choice of parametric and non-parametric tests for paired data should take into account the normality of the data. Therefore, applying the Pearson correlation coefficient (non-parametric test) in non-parametric data increases the chances of spurious associations (by chance or systematic error), which result in a Type I error. Knowing that young researcher and editors of scientific journals might be guided by positive results. It is common for editors to select articles for publication based on  $p < 0.05$  value. However, it would also be important to select papers taking into account the fulfillment of the assumptions for the use of parametric and non-parametric tests. Thus, the aim of the present study was to address the two Pearson and Spearman correlation coefficient tests and to suggest recommendations for practitioners of statistics in the area of Health Sciences for the safe and adequate use of data prior publication.

**Key words:** Correlation; Normality; Bias; Type I error.

**Contato:** Dahan da Cunha Nascimento – dahanc@hotmail.com

Dahan da Cunha Nascimento<sup>1</sup>  
Jeiser Alves de Almeida<sup>2</sup>  
Thiago Hanyel Delgado Carvalho<sup>3</sup>  
Jonato Prestes<sup>4</sup>

<sup>1</sup>Programa de Pós-Graduação Stricto Sensu em Educação Física da Universidade Católica de Brasília.

<sup>2</sup>Pesquisa em Exercício e Nutrição na Saúde e Rendimento Esportivo – PENSARE, Programa de Pós-Graduação em Ciências do Movimento, Universidade Federal de Mato Grosso do Sul.

<sup>3</sup>Curso de Educação Física do Centro Universitário do Distrito Federal.

<sup>4</sup>Programa de Pós-Graduação Stricto Sensu em Educação Física da Universidade Católica de Brasília.

## Introdução

A utilização de procedimentos estatísticos é de fundamental importância para a interpretação apropriada de um conjunto de dados. Trata-se, portanto, de uma das mais importantes ferramentas para compreensão de fenômenos, dentre eles o científico. Os testes estatísticos são, por regra, baseados em pressupostos que podem se diferenciar de acordo com sua especificidade. Contudo, em algumas pesquisas, é comum a prática de associações de causa-efeito, e a partir desta premissa, pode ser observado o viés de aferição (erro sistemático), o qual é influenciado por testes que podem apresentar baixa aderência aos dados analisados de forma intencional ou não-intencional.

Desta forma, a baixa aderência do teste aos dados selecionados pode levar a conclusões inadequadas (associação espúria) como o erro do Tipo I, ou seja, encontrar uma diferença significativa, quando na verdade não existe (acaso, erro aleatório). Por isso, o conhecimento adequado dos pressupostos básicos antes de utilizar determinado teste estatístico, torna-se de fundamental conhecimento.

A escolha do teste paramétrico e não paramétrico para dados pareados deve levar em conta a normalidade dos dados que pode ser avaliada pelos testes de *Shapiro-Wilk* e *Kolmogorov-Smirnov*. Em adendo, os pressupostos utilizados para se utilizar um teste de correlação paramétrica ou não paramétrica, leva em consideração o tipo de variável, que deve ser do tipo quantitativa (contínua ou discreta), deve ser pareado (ex. o índice de massa corporal do participante deve ser pareado com a pressão arterial sistólica do mesmo participante), deve haver normalidade bivariada, ou seja, as duas variáveis devem ser normalmente distribuídas e sem valores discrepantes ou *outliers*, e por fim, deve haver relação linear entre as duas variáveis.

Com isso, testes paramétricos devem ser aplicados em dados paramétricos e testes não-paramétricos devem ser aplicados em dados não-paramétricos. Aplicar o coeficiente de correlação de *Pearson* (teste paramétrico) em dados não paramétricos aumenta as chances de associações espúrias (por acaso ou erro sistemático), as quais resultam em erro do Tipo I, pois apresenta baixo poder para identificar uma diferença genuína quando ela realmente existe. Neste sentido, o valor de  $r$  é facilmente influenciado por um valor discrepante e o coeficiente de correlação de *Spearman* quando aplicado em dados não paramétricos apresenta poder suficiente para identificar se a força de correlação e a direção encontrada entre as duas variáveis é estatisticamente significativa e genuína. O mesmo teste é robusto quando há presença de um valor discrepante ou quando as suposições para a aplicação do coeficiente de correlação de *Pearson* são violadas.

O testes de correlação de *Pearson* e *Spearman* são testes que avaliam a força (ex. fraca, moderada ou forte) e a direção linear entre duas variáveis contínuas<sup>1</sup>. Os valores de correlação

podem variar entre -1 para uma correlação negativa perfeita e +1 para uma correlação positiva perfeita. Quanto mais perto de 0, mais fraca será a correlação entre as duas variáveis<sup>1</sup>.

Outro ponto de destaque que devemos citar é que o teste de correlação não avalia causa-efeito<sup>2</sup>. Portanto, seria incorreto afirmar que o sedentarismo causa hipertensão arterial. Deve-se considerar o fenômeno que ocorre na população, o tipo de associação bem como o modelo causal. Assim, considera-se que o sedentarismo explica parte da variação sobre a doença, e obviamente outras inúmeras variáveis não controladas (ex. consumo de cigarros) interferem no resultado final. Portanto, estudos com hipóteses correlacionais não são delineados originalmente para responder a pergunta, mas sim para indicar uma direção de pesquisa de futuros estudos.

Tendo uma visão preliminar do que foi abordado até o momento, podemos entender que a correlação de *Pearson* quando aplicada em um dado não normal, induzirá ao erro do tipo I (encontrar um diferença quando na verdade ela não existe), enquanto que a correlação de *Spearman* não rejeitará a hipótese nula. Entendendo que na grande maioria das vezes o pensamento do jovem pesquisador e também de editores de periódicos científicos serão guiados por resultados positivos, ou seja, aceitando a hipótese alternativa e rejeitando a hipótese nula. É comum a possibilidade de editores selecionarem artigos para publicação tendo como base o valor de  $p < 0,05$ . Contudo, o também seria importante selecionar os artigos levando em consideração os cumprimentos dos pressupostos para a utilização de testes paramétricos e não-paramétricos.

Ao tomar conhecimento do que foi abordado, a ciência depende da exposição completa da forma de avaliação dos dados. No entanto, parece existir uma padronização de como o pesquisador decide sobre o que realmente vale a pena reportar e a decisão de reportar os achados (se foram significativos ou não) se torna subjetiva<sup>3</sup> e as vezes intencional. Isso pode levar ao viés de publicação, ou seja, quando os resultados da publicação são baseados na direção ou achados significativos, quando na verdade não deveriam ser<sup>3</sup>.

Portanto, o objetivo do presente estudo foi abordar os dois testes de coeficiente de correlação de *Pearson* e *Spearman* e sugerir recomendações para praticantes de estatística na área de Ciências da Saúde para a utilização segura e adequada dos dados antes da publicação.

## Método

Para o melhor entendimento do leitor e demonstração dos principais objetivos do estudo, foi utilizado no presente artigo dados originais de um estudo transversal realizado com mulheres idosas hipertensas e normotensas. O estudo foi aprovado pelo Comitê de Ética e Pesquisa local (protocolo: 45648115.8.0000.5650/2016), cada participante foi informada dos benefícios e riscos do estudo bem como da assinatura do termo de consentimento livre esclarecido foi assinado.

## Participantes

Noventa idosas selecionadas por conveniência ( $68,13 \pm 6,02$  anos;  $1,54 \pm 0,06$  m;  $68,55 \pm 11,32$  kg;  $28,83 \pm 4,39$  kg/m<sup>2</sup>) participaram do estudo. Todas as idosas foram entrevistadas e responderam a anamnese sobre ao nível de atividade física e uso de medicamentos. Em um outro momento, a composição corporal foi avaliada pelo método de densitometria por dupla emissão de raios-X (DEXA) e logo após foram submetidas ao teste de esforço incremental em rampa, seguindo o protocolo adotado por Silva et al. (2017)<sup>4</sup>. A partir do teste de esforço, informações como equivalente metabólico (METs), recuperação de frequência cardíaca durante 1 minuto após o teste, índice de reserva cronotrópico e uso de betabloqueadores e/ou inibidores de canais de cálcio foram utilizadas para o cálculo da idade fisiológica baseada no teste de esforço (IBTE), a qual recentemente foi considerada como melhor preditora para todas as causas de mortalidade quando comparado com a idade cronológica<sup>5</sup>. A partir disso, a força de correlação e a direção linear entre as variáveis IBTE e pressão arterial sistólica (PAS) foram examinadas. A PAS representa um importante preditor para morbidade cardiovascular e mortalidade<sup>6-8</sup>.

Os softwares *Statistical Package for Social Science* (SPSS 18.0, Chicago, IL) e *GraphPad Prism 7* (San Diego, CA) foram utilizados para as análises estatísticas e o nível de significância adotado foi de  $p \leq 0,05$ .

## Correlação da IBTE com a pressão arterial sistólica usando diferentes testes estatísticos

As hipóteses correlacionais do presente estudo são citados abaixo:

- Hipótese nula ( $H_0$ ) = Não haverá correlação significativa entre IBTE e a pressão arterial sistólica (PAS).
- Hipótese alternativa ( $H_1$ ) = Haverá correlação significativa entre IBTE e a PAS.

Inicialmente para saber qual teste deverá ser aplicado, o teste de normalidade de *Shapiro-Wilk* deve ser realizado. Recomendamos nesse artigo o teste de *Shapiro-Wilk*, pois o mesmo é superior e apresenta poder acima de 80% quando comparado com o teste *Kolmogorovi-Smirnov* para uma amostra de 90 participantes<sup>9</sup>. Pois, possui menor risco de cometer o erro do Tipo I, ou seja, reportar que os dados são normalmente distribuídos quando na verdade não são.

## Resultados

Ao analisar o quadro 1, verificamos no *output* que não houve normalidade bivariada, ou seja, somente a variável PAS apresenta distribuição normal ( $p > 0,05$ ) enquanto que a variável

IBTE não ( $p < 0,05$ ). Com isso, a suposição para o teste paramétrico de *Pearson* foi violada e o teste mais indicado para a avaliação da correlação entre as duas variáveis é o teste de *Spearman*.

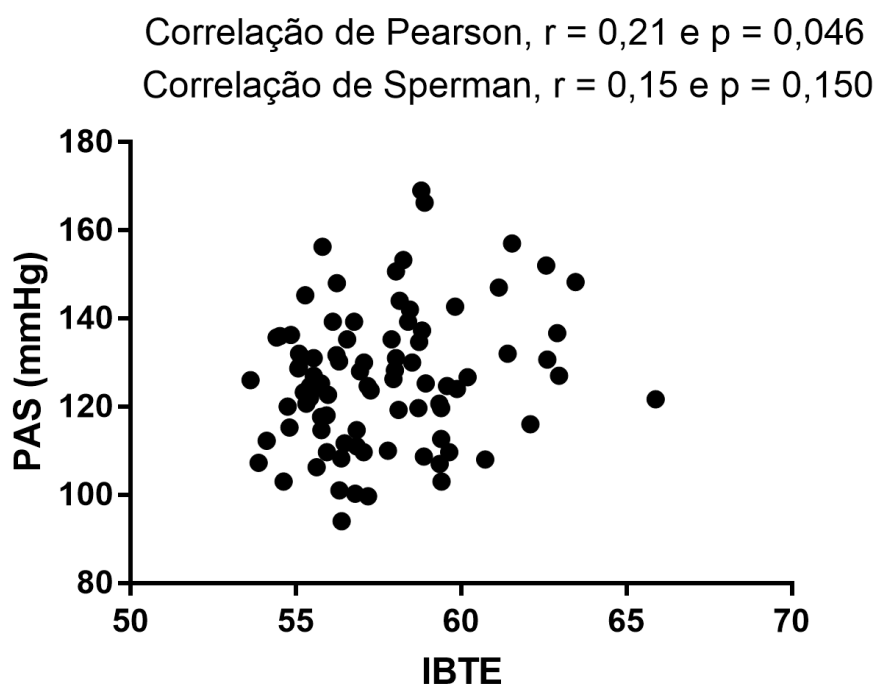
**Quadro 01.** Testes de normalidade e variáveis de investigação.

	<i>Kolmogorov-Smirnov</i>			<i>Shapiro-Wilk</i>		
	Statistic	df	Sig.	Statistic	df	Sig.
PAS (mmHg)	,076	90	,200	,984	90	,324
IBTE	,106	90	,014	,939	90	,000

Abreviação: PAS = pressão arterial sistólica; IBTE = idade fisiológica baseada no teste de esforço; df = *degrees of freedom*; sig = *significant*.

Observando a figura 1, nota-se que usando diferentes coeficientes de correlação, o erro do Tipo I pode vir acontecer. Ou seja, ao interpretar o coeficiente de correlação de *Pearson*, reportamos que a correlação entre as variáveis IBTE e PAS foi significativa e positiva, mas se a correlação de *Spearman* for levada em consideração, não rejeita-se a hipótese nula ( $H_0$ ). Portanto, considerando o rigor na hora da publicação e principalmente na análise dos dados, o pesquisador deve selecionar o teste mais apropriado aos dados e não aquele que deve corroborar com a hipótese alternativa ( $H_1$ ). O pesquisador nesse momento deve evitar o viés de publicação, ou seja, aceitar que não há diferença, que as suposições foram violadas e reportar de forma ética e imparcial os resultados do estudo<sup>3</sup>.

**Figura 1.** Correlação entre IBTE e PAS usando diferentes métodos.



Abreviação: PAS = pressão arterial sistólica; IBTE = idade fisiológica baseada no teste de esforço.

Ao mesmo tempo, o pesquisador pode escolher corrigir os dados ou transformá-los antes de aplicar o teste não paramétrico. Neste sentido, a correção dos dados e sua transformação poderia alterar conclusão do estudo? Tem sido recomendado que o tratamento dos dados seja realizada quando não apresentarem normalidade e caso falhe, considera-se a substituição do valor<sup>1</sup>. Assim, por meio do processo teórico-prático, o pesquisador pode utilizar com cautela as dicas elencadas no Quadro 2 que foi criada para melhor orientar o praticante de estatística. No entanto, sabemos que na seção estatística de um artigo científico o passo a passo da tabela 2 dificilmente é reportado.

**Quadro 2.** Passo a passo para a decisão do tipo de teste de coeficiente de correlação que será aplicado durante a análise dos dados.

<b>PASSO 1</b>		
<b>Verificar se os seus dados possuem <i>outliers</i> pelos métodos abaixo</b> <i>Box-Plot.</i> Escores-z.		
<b>Sim, os meus dados possuem <i>outliers</i></b> Vá para o passo 2.	<b>Não, os meus dados não possuem <i>outliers</i></b> Vá para o passo 3.	
<b>PASSO 2</b>		
<b>Quer remover o caso?</b> Sim (voltar para o passo 1); Não (vá para o passo 3).	<b>Quer transformar os dados?</b> Sim (voltar para o passo 1) Não (vá para o passo 3).	<b>Quer substituir o valor?</b> Sim (voltar para o passo 1) Não (vá para o passo 3).
<b>PASSO 3</b>		
<b>Utilize um teste objetivo para avaliar a normalidade dos dados.</b> <i>Shapiro-Wilk.</i>		
<b>Após realizar o teste de normalidade, os dados apresentam normalidade?</b> Sim (Vá para o passo 4). Não (Vá para o passo 5).		
<b>PASSO 4</b>		
Aplique o teste paramétrico coeficiente de correlação de <i>Pearson</i> .		
<b>PASSO 5</b>		
Aplique o teste não paramétrico coeficiente de correlação de <i>Spearman</i> .		

## Discussão

Parafraseando a citação do pesquisador Dickersin (1999)<sup>3</sup>, se o viés de publicação existe, será que é importante se preocupar? Primeiramente vamos lembrar que o viés de publicação acontece quando os resultados da publicação são baseados na direção ou achados significativos<sup>3</sup>. Com isso, baseado nos resultados da presente reflexão, *Pearson* ou *Spearman*, eis a questão?

Os resultados do estudo somente devem ser divulgados quando houver diferença significativa? O que tem sido feito para investigar a possibilidade do viés de publicação na grande área das Ciências da Saúde?

É notório o esforço e motivação que os pesquisadores exercem em oferecer resultados positivos para publicação ao invés de não rejeitar a hipótese nula. Adicionalmente, há ainda editores de periódicos especializados que selecionam estudos tendo como base o valor de  $p$ , esperando que este apresente significância, ou seja, abaixo de  $0,05^{10}$ .

Independentemente da subárea de atuação do profissional da Ciências da Saúde, as decisões em sua prática clínica deve, sempre que possível, ser pautada à luz da evidência científica. Logo, enquanto leitores e consumidores de pesquisa científica, tanto os pesquisadores quanto o corpo editorial dos periódicos devem estar atentos as publicações que remetam também a hipótese nula e não apenas a hipótese alternativa.

## Conclusão

Testes paramétricos devem ser aplicados em dados normalmente distribuídos que satisfaçam as suposições básicas. A aplicação em dados não distribuídos em normalidade levará ao erro do Tipo I, pois o teste não apresenta robustez suficiente aos valores discrepantes. A aplicação correta de testes estatísticos e com rigor diminuirá a ocorrência do viés de publicação e consequentemente aumentará a qualidade das intervenções práticas. Afinal, o que é a teoria sem a prática e a prática sem a estatística?

## Conflitos de Interesse

Os autores declaram não haver conflito de interesse.

## Agradecimentos

O primeiro autor dedica esse trabalho a sua família (Rita Cunha e Nicolas Cunha).

## Referências

1. Field A. Discovering statistics using IBM SPSS statistics: sage; 2013.
2. Dankel SJ, Loenneke JP. Effect Sizes for Paired Data Should Use the Change Score Variability Rather Than the Pre-test Variability. *Journal of strength and conditioning research* 2018.
3. Dickersin K. The existence of publication bias and risk factors for its occurrence. *Jama* 1990; 263(10): 1385-9.
4. Silva CRd, Saraiva B, Nascimento DC, et al. Relationship between adiposity and heart rate recovery following an exercise stress test in obese older women. *Revista Brasileira de Cineantropometria & Desempenho Humano* 2017; 19(5): 554-64.
5. Harb SC, Cremer PC, Wu Y, et al. Estimated age based on exercise stress testing performance outperforms chronological age in predicting mortality. *Eur J Prev Cardiol* 2019: 2047487319826400.
6. Lewington S, Clarke R, Qizilbash N, Peto R, Collins R, Prospective Studies C. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet* 2002; 360(9349): 1903-13.
7. Lawes CM, Bennett DA, Feigin VL, Rodgers A. Blood pressure and stroke: an overview of published reviews. *Stroke* 2004; 35(4): 1024.
8. Wang JG, Staessen JA, Franklin SS, Fagard R, Gueyffier F. Systolic and diastolic blood pressure lowering as determinants of cardiovascular outcome. *Hypertension* 2005; 45(5): 907-13.
9. da Cunha Nascimento D, Tibana RA, de Melo GF, Prestes J. Testes de normalidade em análises estatísticas: uma orientação para praticantes em ciências da saúde e atividade física. *Revista Mackenzie de Educação Física e Esporte* 2017; 14(2).
10. Maxwell C. Clinical trials, reviews, and the Journal of Negative Results. *British journal of clinical pharmacology* 1981; 11(1): 15-8.