

# Análise de correlação em estudos clínicos e experimentais

## Correlation analysis in clinical and experimental studies

Hélio Amante Miot<sup>1</sup>

É habitual que, em uma pesquisa clínica ou biomédica, seja do interesse do pesquisador investigar se os valores de duas ou mais variáveis quantitativas se modificam de forma conjunta em um mesmo sujeito ou objeto de estudo. Ou seja, quando o valor de uma variável aumenta, o valor de outra tende a aumentar; ou, inversamente, reduza-se progressivamente. Há uma série de testes estatísticos que exploram a intensidade e o sentido desse comportamento mútuo entre variáveis, os chamados testes de correlação<sup>1,2</sup>.

O primeiro passo para analisar a correlação entre duas variáveis quantitativas deve ser a visualização do diagrama de dispersão, a fim de identificar se existe uma variabilidade gradual entre os conjuntos de dados, se essa variação é monotônica (predominantemente ascendente ou descendente), se assume uma tendência proporcional (linear) e se a distribuição subjacente dos dados é normal<sup>2-4</sup>. Diferentes combinações dessas premissas indicam diferentes técnicas de análise de correlação.

A Figura 1 exemplifica a dispersão dos valores entre quatro variáveis hipotéticas (V1, V2, V3 e V4), as quais apresentam dados aderentes à distribuição normal (Shapiro-Wilk,  $p > 0,32$ ).

As variáveis V1 e V2 apresentam um crescimento simultâneo dos valores, que se distribuem sobre uma linha reta subjacente imaginária (ideal), a qual descreve a trajetória dos dados. Pode-se afirmar que existe correlação linear positiva entre V1 e V2. Da mesma forma, Rossi e colaboradores identificaram forte correlação positiva ( $\rho = 0,82$ ;  $p < 0,01$ ) entre os escores da Escala de Gravidade Clínica dos Sintomas Venosos e da escala de dor em doença venosa crônica<sup>5</sup>.

As variáveis V1 e V3, em contrapartida, apresentam um comportamento antagônico: enquanto os valores de uma aumentam, os da outra reduzem. Pode-se afirmar que existe correlação linear negativa entre V1 e V3, assim como Ohki e Bellen identificaram moderada correlação negativa ( $\rho = -0,65$ ;  $p < 0,01$ )

entre a temperatura regional média e a incidência de trombose venosa<sup>6</sup>.

Nota-se ainda que os valores da correlação V1 e V3 aproximam-se mais da linha reta imaginária que os valores da correlação entre V1 e V2. Isso leva à conclusão que a relação entre os valores das variáveis V1 e V3 é mais intensa que a relação entre V1 e V2, apesar dos sentidos serem opostos.

As comparações dos dados da V4, sejam com V1, V2 ou V3, não apresentam um comportamento gradualmente ascendente ou descendente. Isso leva à conclusão que V4 não apresenta correlação com as demais variáveis.

A técnica mais difundida para a avaliação da correlação entre duas variáveis quantitativas é o coeficiente de correlação produto-momento de Pearson, ou  $r$  de Pearson, que pressupõe distribuição normal das duas amostras e comportamento linear da relação entre as variáveis<sup>2,7</sup>. A falta de observação dessas premissas leva a conclusões equivocadas, mesmo diante de amostragem numerosa.

É, porém, bastante comum que amostras de dados clínicos e demográficos não sigam distribuição normal (por exemplo, distribuição de renda, índices de qualidade de vida, índices de gravidade de uma doença, anos de estudo, número de filhos). As opções mais utilizadas na investigação da correlação entre variáveis que não apresentam distribuição normal são a correlação da ordem de postos de Spearman e o coeficiente de correlação de postos de Kendall ( $Tau-b$ ), que substituem os dados originais por postos ordenados ( $ranks$ )<sup>2,7,8</sup>. Esses métodos também se impõem em casos que ao menos uma das variáveis apresenta características ordinais (por exemplo, classe funcional, nível de escolaridade, estadiamento de câncer, classe social).

Outra vantagem do uso dos testes não paramétricos de Spearman e de Kendall é que eles não se restringem a correlações lineares, desde que apresentem

<sup>1</sup> Universidade Estadual Paulista – UNESP, Faculdade de Medicina de Botucatu, Departamento de Dermatologia e Radioterapia, Botucatu, SP, Brasil.

Fonte de financiamento: Nenhuma.

Conflito de interesse: O autor declara não haver conflitos de interesse que precisam ser informados.

Submetido em: Fevereiro 09, 2018. Aceito em: Fevereiro 12, 2018.

O estudo foi realizado no Departamento de Dermatologia e Radioterapia, Faculdade de Medicina de Botucatu, Universidade Estadual Paulista (UNESP), Botucatu, SP, Brasil.

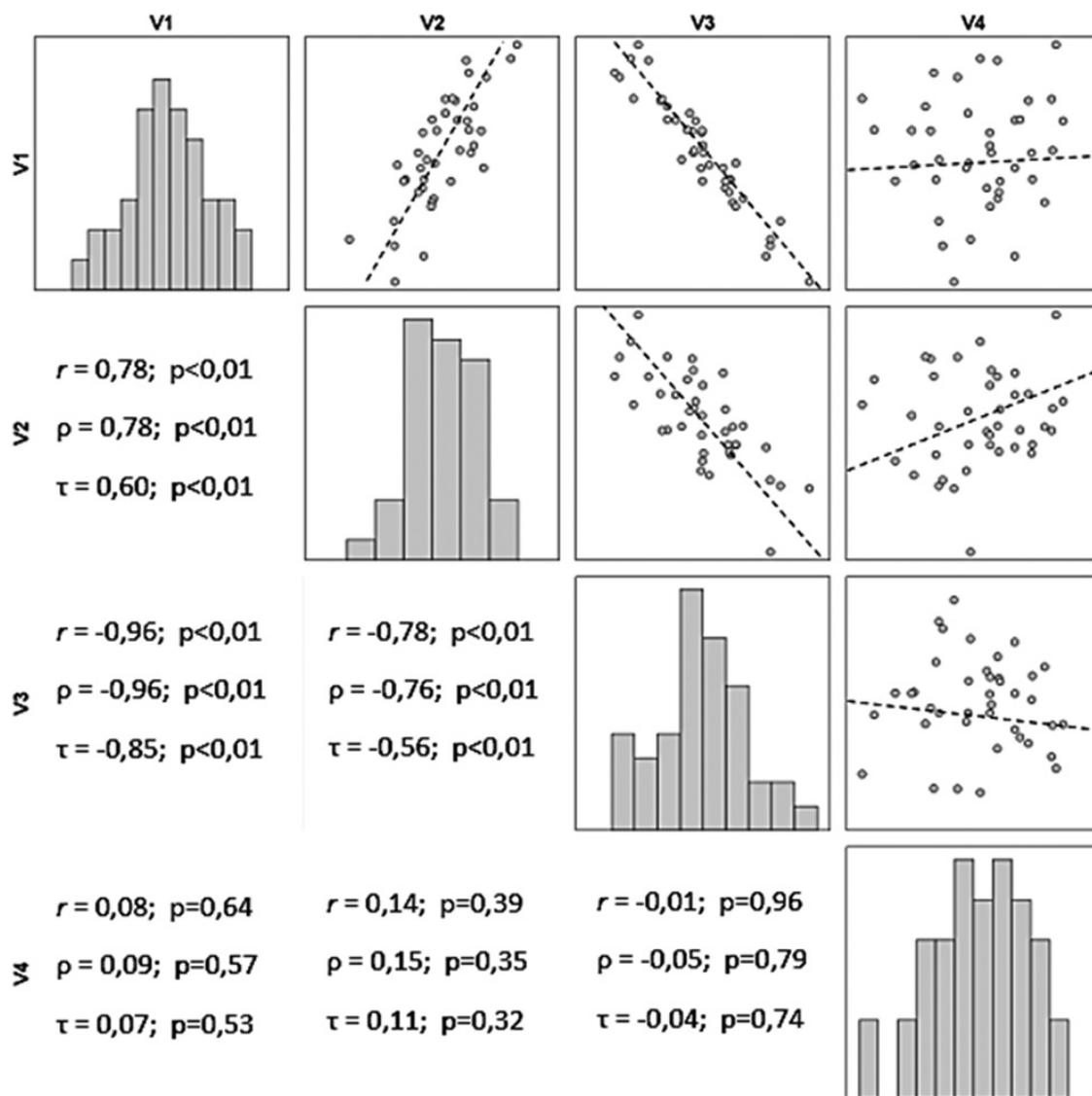


Figura 1. Histogramas, diagramas de dispersão e coeficientes de correlação ( $r$ : Pearson,  $\rho$ : Spearman, e  $\tau$ : Kendall *Tau*-b) entre quatro variáveis quantitativas hipotéticas V1, V2, V3 e V4 ( $n = 40$ ).

comportamento monotônico. Ou seja, eles devem apresentar uma relação gradual no mesmo sentido (ascendente ou descendente) para todo o domínio de dados estudado.

Na Figura 2, observa-se que não ocorre proporcionalidade direta (linear) entre a variação dos dados de V1 e V5, mas um incremento aparentemente exponencial. Como a variação é monotônica (os dados de V1 oscilam em um crescente em função de V5), os coeficientes de Spearman e Kendall podem estimar a correlação. Já nesse exemplo, para o emprego do coeficiente de Pearson, seria necessária a transformação logarítmica dos dados para garantir certa linearidade da correlação (Figura 2: V1 x V6). Destaque-se que

os coeficientes  $\rho$  e  $\tau$  resultaram nos mesmos valores para a correlação V1 *versus* V5 e para V1 *versus* V6, já que V6 se tratava de uma transformação de V5 em dados monotônicos.

O coeficiente de Spearman ( $\rho$  ou rho) é o mais empregado em ciências biomédicas para avaliar a correlação entre duas variáveis quantitativas, provavelmente porque seja semelhante ao método de Pearson, desde que os dados sejam substituídos por postos ordenados (*ranks*). O que, porém, inspira cuidado na generalização das conclusões quanto à interpretação da correlação entre os valores dos *ranks* dos dados das variáveis e não dos dados originais.

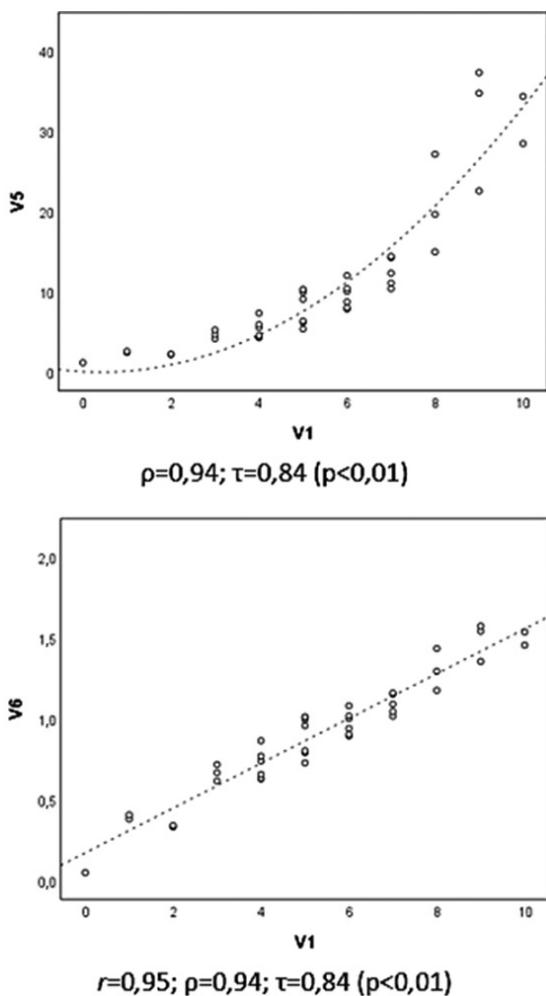


Figura 2. Diagramas de dispersão e coeficientes de correlação entre as variáveis hipotéticas V1, V5 e V6 ( $n = 40$ ), sendo V6 a transformação logarítmica de V5 ( $V6 = \log_{10} V5$ ).

O coeficiente de Kendall Tau-b ( $\tau$  ou  $t_b$ ), em contrapartida, apresenta propriedades matemáticas que o tornam mais robusto a dados extremos (*outliers*), de maior capacidade de inferência populacional e menor erro de estimativa. Apesar da significância (p-valor) e o sentido (+ ou -) se aproximarem do método de Spearman, seu coeficiente apresenta valores menos proeminentes e sua interpretação é diferenciada, significando o percentual de pares observados que assumem o mesmo sentido na amostra (concordantes) menos os pares discordantes. Por exemplo, um coeficiente  $\tau$  de 0,60 representa que 80% dos pares são concordantes, enquanto 20% são discordantes ( $\tau = 0,80 - 0,20 = 0,60$ )<sup>9</sup>.

A transformação dos dados (por exemplo, logaritmo, raiz quadrada,  $1/x$ ) a fim de obter sua normalização para testar o coeficiente de Pearson é

uma alternativa válida em amostras com distribuição dos dados assimétricas (Figura 2: V1 x V6). Entretanto, deve-se ter em mente que, da mesma forma que as técnicas que empregam postos ordenados (*ranks*), as transformações dos dados alteram a escala entre as medidas, influenciando a interpretação direta das medidas de efeito<sup>7</sup>.

A magnitude do efeito da correlação entre duas ou mais variáveis é representada pelos coeficientes de correlação, que assumem valores de -1 até +1, passando pelo zero (ausência de correlação). Coeficientes positivos ( $r > 0$ ) indicam correlação direta (Figura 1: V1 x V2) entre as variáveis; já coeficientes negativos ( $r < 0$ ) significam uma correlação inversa (Figura 1: V1 x V3 e V2 x V3).

Cada teste de correlação apresenta um coeficiente individualizado, que demanda uma interpretação própria. De forma geral, para os coeficientes  $r$  de Pearson e  $\rho$  de Spearman, valores entre 0 e 0,3 (ou 0 e -0,3) são biologicamente desprezíveis; entre 0,31 e 0,5 (ou -0,31 e -0,5) são correlações fracas; entre 0,51 e 0,7 (ou -0,51 e -0,7) são moderadas; entre 0,71 e 0,9 (ou -0,71 e -0,9) são correlações fortes; e  $> 0,9$  (ou  $< -0,9$ ) são consideradas muito fortes<sup>8</sup>.

Uma particularidade do coeficiente  $r$  de Pearson é que o quadrado do seu valor representa uma estimativa do percentual de variabilidade dos valores de uma variável que é explicado pela variabilidade da outra. Por exemplo, um coeficiente  $r = 0,7$  significa que 49% da variabilidade de uma variável pode ser explicada, ou é acompanhada pela variação dos valores da outra na amostra testada.

Em estudos clínicos e biomédicos, a maior parte dos coeficientes com significado biológico situa-se entre 0,5 e 0,8 (ou -0,5 e -0,8). Isso decorre tanto da existência de erros de mensuração, técnicas laboratoriais ou variação de instrumentos, que interferem na exatidão e na precisão das medidas, mas, principalmente, porque fenômenos biológicos sofrem influências multifatoriais e interações complexas, em que a variação de apenas uma variável não explica totalmente o comportamento de outra<sup>2</sup>.

Os testes de significância da correlação entre variáveis quantitativas se baseiam na hipótese nula que não existe nenhuma correlação entre as variáveis ( $r = 0$ ), sujeitando o p-valor à influência tanto da dimensão do efeito quanto do tamanho amostral. Isso exige cautela na interpretação de coeficientes que resultem em baixa correlação ( $r < 0,3$ ), mas que apresentem p-valores muito significativos, decorrentes de amostras superdimensionadas. O cálculo do tamanho amostral para análise de correlação já foi explorado anteriormente nesta revista<sup>10</sup>.

Os coeficientes de correlação apresentam propriedades inferenciais e devem ser preferencialmente expressos, em textos científicos, pelo seu intervalo de confiança de 95% e a significância (p-valor), por exemplo:  $\rho = 0,76$  (IC95% 0,61-0,91),  $p < 0,01$ <sup>11,12</sup>. No caso de múltiplas comparações, o coeficiente pode ser apresentado de forma isolada, na forma de uma matriz, e com sua significância sinalizada para facilitar a leitura dos dados, como Brianezi e colaboradores representaram as 28 correlações de sete parâmetros histológicos em uma tabela<sup>13</sup>. Casos especiais que envolvam centenas ou milhares de correlações podem exigir técnicas de representação gráfica, como os mapas térmicos de cor (*heatmap*) frequentemente usados em estudos genômicos, como Hsu e colaboradores representaram 4.930 correlações entre (85x58) variáveis genômicas e metabolômicas<sup>14</sup>.

Em casos de dados ordinais com poucas categorias (por exemplo, escores de satisfação, itens de qualidade de vida, nível socioeconômico), a exploração baseada no teste de correlação policórica pode ser mais robusta (menor erro tipo I) que os testes de Spearman e Kendall<sup>15</sup>. Ainda, apesar de pouco utilizadas, há também formas de se avaliar a correlação entre variáveis de natureza categóricas (por exemplo, coeficiente V de Cramér) e entre variáveis dicotômicas e quantitativas (por exemplo, coeficiente de correlação ponto-biserial), todavia, essas abordagens ultrapassam o escopo deste texto<sup>7</sup>.

Em situações especiais em que a correlação linear entre diferentes variáveis deva ser analisada conjuntamente (por exemplo, itens de um questionário) visando a compreensão da variação conjunta global das variáveis, a análise de correlação entre “n” variáveis pode ser avaliada pelo coeficiente de correlação intraclasse (CCI) do tipo consistência. Há diferentes formas de análise do CCI, que levam a indicadores de diferentes magnitudes<sup>16</sup>. O CCI (aleatório ou misto de dois fatores, medidas médias) retorna o mesmo valor do coeficiente  $\alpha$  de Cronbach, usado para medir a consistência interna de escalas<sup>17</sup>.

A identificação de uma correlação significativa entre duas ou mais variáveis deve ser interpretada com cautela, visto que a análise estatística não fornece evidências de dependência direta ou mesmo de causalidade entre as variáveis, mas apenas que elas tendem a variar conjuntamente<sup>1,18,19</sup>. Entretanto, apesar do risco de fálacia ao se concluir causalidade a partir de resultados de correlação entre variáveis, os testes de correlação são importantes técnicas exploratórias para a investigação de associação entre o comportamento de grupos de variáveis, favorecendo a elaboração de modelos hipotéticos que devam ser confirmados posteriormente por meio de experimentos dedicados. Aliás, isso ocorre com estudos clínicos ecológicos, que frequentemente empregam técnicas de correlação na análise de seus dados e servem

de base para investigação posterior dos fenômenos indicados pela correlação entre indicadores e grupos populacionais<sup>20-22</sup>. Ou ainda estudos tipo *genome-wide* e de proteômica, ditos exploratórios, que estudam os padrões de correlações dos achados com as variáveis clínicas em busca da indicação de modelos a serem comprovados posteriormente<sup>19,23</sup>.

Inclusive, quando se realizam múltiplos testes de correlação em uma sequência de variáveis, aumentam-se as chances de se identificar, ao acaso, correlações ditas “espúrias”, que devem ser avaliadas quanto à plausibilidade biológica e comprovadas posteriormente por técnicas adequadas de investigação. O emprego de técnicas de correção dos p-valores ajustadas para as correlações múltiplas é sempre indicado nessas condições<sup>7,19,24-27</sup>.

Outra limitação de natureza inferencial das análises de correlação baseia-se na incapacidade de extrapolação das conclusões para outros intervalos de dados ou para populações diferentes dos estudados.

Análises de correlação não foram desenvolvidas *a priori* com a finalidade de predição de valores ou da inferência da participação de múltiplas variáveis na explicação de um fenômeno, havendo técnicas de regressão ou de análise multivariadas com essas finalidades<sup>7</sup>. Apesar de existirem técnicas de correlação parcial que ajustam os valores da correlação para o comportamento de variáveis de confusão (idêntico ao coeficiente  $\beta$  padronizado em regressão linear múltipla), assim como técnicas de transformação polinomial para corrigir correlações não monotônicas, um profissional estatístico experiente deve ser consultado para o planejamento e a execução de análises de maior complexidade.

Análises de correlação podem também ser empregadas na comparação do paralelismo das medidas entre duas escalas diferentes para a medida do mesmo fenômeno, como escalas psicométricas de qualidade de vida<sup>28</sup>, ou clinimétricas, como escalas de risco para úlcera por pressão. Contudo, muitas vezes, elas são utilizadas por pesquisadores no sentido equivocado de se testar a concordância dos dados ou medidas consecutivas do mesmo fenômeno (por exemplo, teste-reteste<sup>29</sup>, calibragem de instrumentos de medida, comparação entre juízes), enquanto existem métodos mais adequados para essas finalidades<sup>30</sup>.

Finalmente, as estratégias de avaliação da correlação entre variáveis devem ser incentivadas em pesquisa clínica e biomédica, uma vez que maximizam a compreensão do fenômeno estudado. Entretanto, devido às particularidades inerentes aos seus diversos métodos, elas devem ser detalhadamente descritas na metodologia e na apresentação dos resultados.

## REFERÊNCIAS

- Liu J, Tang W, Chen G, Lu Y, Feng C, Tu XM. Correlation and agreement: overview and clarification of competing concepts and measures. *Shanghai Arch Psychiatry*. 2016;28(2):115-20. PMID:27605869.
- Zou KH, Tuncali K, Silverman SG. Correlation and simple linear regression. *Radiol*. 2003;227(3):617-22. <http://dx.doi.org/10.1148/radiol.2273011499>. PMID:12773666.
- Scott M, Flaherty D, Currall J. Statistics: are we related? *J Small Anim Pract*. 2013;54(3):124-8. <http://dx.doi.org/10.1111/jsap.12029>. PMID:23458641.
- Miot HA. Assessing normality of data in clinical and experimental trials. *J Vasc Bras*. 2017;16:88-91. <http://dx.doi.org/10.1590/1677-5449.041117>. PMID:29930631.
- Rossi FH, Volpato MG, Metzger PB, et al. Relationships between severity of signs and symptoms and quality of life in patients with chronic venous disease. *J Vasc Bras*. 2015;14(1):22-8. <http://dx.doi.org/10.1590/1677-5449.20140039>.
- Ohki AV, van Bellen B. Regional incidence of venous thromboembolism in Brazil. *J Vasc Bras*. 2017;16:227-31. <http://dx.doi.org/10.1590/1677-5449.000517>. PMID:29930651.
- Norman GR, Streiner DL. *Biostatistics. The bare essentials*. 4th ed. Shelton: People's Medical Publishing House; 2014.
- Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J*. 2012;24(3):69-71. PMID:23638278.
- Arndt S, Turvey C, Andreasen NC. Correlating and predicting psychiatric symptom ratings: Spearman's r versus Kendall's tau correlation. *J Psychiatr Res*. 1999;33(2):97-104. [http://dx.doi.org/10.1016/S0022-3956\(98\)90046-2](http://dx.doi.org/10.1016/S0022-3956(98)90046-2). PMID:10221741.
- Miot HA. Sample size in clinical and experimental trials. *J Vasc Bras*. 2011;10:275-8. <http://dx.doi.org/10.1590/S1677-54492011000400001>.
- Kim N, Fischer AH, Dyring-Andersen B, Rosner B, Okoye GA. *Research Techniques Made Simple: Choosing Appropriate Statistical Methods for Clinical Research*. *J Invest Dermatol*. 2017;137(10):e173-8. <http://dx.doi.org/10.1016/j.jid.2017.08.007>. PMID:28941476.
- Curran-Everett D, Benos DJ. Guidelines for reporting statistics in journals published by the American Physiological Society. *Am J Physiol Endocrinol Metab*. 2004;287(2):E189-91. <http://dx.doi.org/10.1152/ajpendo.00213.2004>. PMID:15271643.
- Brianezi G, Minicucci EM, Marques ME, Miot HA. Evaluation epidermal p53 immunostaining by digital image analysis. *Skin Res Technol*. 2013;19(1):e108-12. <http://dx.doi.org/10.1111/j.1600-0846.2012.00616.x>. PMID:22672779.
- Hsu HH, Araki M, Mochizuki M, et al. A Systematic Approach to Time-series Metabolite Profiling and RNA-seq Analysis of Chinese Hamster Ovary Cell Culture. *Sci Rep*. 2017;7(1):43518. <http://dx.doi.org/10.1038/srep43518>. PMID:28252038.
- Jin S, Yang-Wallentin F. Asymptotic Robustness Study of the Polychoric Correlation Estimation. *Psychometrika*. 2017;82(1):67-85. <http://dx.doi.org/10.1007/s11336-016-9512-2>. PMID:27660261.
- Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420-8. <http://dx.doi.org/10.1037/0033-2909.86.2.420>. PMID:18839484.
- Sijtsma K. On the use, the misuse, and the very limited usefulness of Cronbach's Alpha. *Psychometrika*. 2009;74(1):107-20. <http://dx.doi.org/10.1007/s11336-008-9101-0>. PMID:20037639.
- Jupiter DC. How are we related? Causality, correlation, and association. *J Foot Ankle Surg*. 2012;51(5):702-3. <http://dx.doi.org/10.1053/j.jfas.2012.05.025>. PMID:22766190.
- Ness RO, Sachs K, Vitek O. From correlation to causality: statistical approaches to learning regulatory relationships in large-scale biomolecular investigations. *J Proteome Res*. 2016;15(3):683-90. <http://dx.doi.org/10.1021/acs.jproteome.5b00911>. PMID:26731284.
- Ojha RP, Offutt-Powell TN, Evans EL, Singh KP. Correlation coefficients in ecologic studies of environment and cancer. *Arch Environ Occup Health*. 2011;66(4):241-4. <http://dx.doi.org/10.1080/19338244.2010.539641>. PMID:22014198.
- Schmitt JV, Chinem VP, Marques MEA, Miot HA. Increase in the incidence of basal cell carcinoma in a university hospital between 1999 and 2009. *An Bras Dermatol*. 2011;86(2):375-7. <http://dx.doi.org/10.1590/S0365-05962011000200029>. PMID:21603829.
- Schmitt JV, Miot HA. Distribution of Brazilian dermatologists according to geographic location, population and HDI of municipalities: an ecological study. *An Bras Dermatol*. 2014;89(6):1013-5. <http://dx.doi.org/10.1590/abd1806-4841.20143276>. PMID:25387516.
- Kumari S, Nie J, Chen HS, et al. Evaluation of gene association methods for coexpression network construction and biological knowledge discovery. *PLoS One*. 2012;7(11):e50411. <http://dx.doi.org/10.1371/journal.pone.0050411>. PMID:23226279.
- Boardman TJ, Moffitt DR. Graphical Monte Carlo type I error rates for multiple comparison procedures. *Biometrics*. 1971;27(3):738-44. <http://dx.doi.org/10.2307/2528613>. PMID:5116579.
- Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. *Stat Med*. 1990;9(7):811-8. <http://dx.doi.org/10.1002/sim.4780090710>. PMID:2218183.
- Miot HA, Brianezi G. Morphometric analysis of dermal collagen by color clusters segmentation. *An Bras Dermatol*. 2010;85(3):361-4. <http://dx.doi.org/10.1590/S0365-05962010000300010>. PMID:20676470.
- Ludbrook J. Multiple comparison procedures updated. *Clin Exp Pharmacol Physiol*. 1998;25(12):1032-7. <http://dx.doi.org/10.1111/j.1440-1681.1998.tb02179.x>. PMID:9888002.
- Maranzatto CF, Miot HA, Miot LD, Meneguim S. Psychometric analysis and dimensional structure of the Brazilian version of melasma quality of life scale (MELASQoL-BP). *An Bras Dermatol*. 2016;91(4):422-8. <http://dx.doi.org/10.1590/abd1806-4841.20165014>. PMID:27579735.
- Paula HR, Haddad A, Weiss MA, Dini GM, Ferreira LM. Translation, cultural adaptation, and validation of the American Skindex-29 quality of life index. *An Bras Dermatol*. 2014;89(4):600-7. <http://dx.doi.org/10.1590/abd1806-4841.20142453>. PMID:25054747.
- Miot HA. Agreement analysis in clinical and experimental trials. *J Vasc Bras*. 2016;15:89-92. <http://dx.doi.org/10.1590/1677-5449.004216>. PMID:29930571.

### Correspondência

Hélio Amante Miot  
 Universidade Estadual Paulista – UNESP  
 Av. Prof. Mário Rubens Guimarães Montenegro, s/n - Distrito de  
 Rubião Junior  
 CEP 18618-687 - Botucatu (SP), Brasil  
 Tel.: (14) 3882-4922  
 E-mail: heliomiot@gmail.com

### Informações sobre o autor

HAM - Livre-docente, Departamento de Dermatologia e Radioterapia, Faculdade de Medicina de Botucatu, Universidade Estadual Paulista (UNESP).