

CRITÉRIOS DE AVALIAÇÃO DA QUALIDADE DE RESULTADOS DE PESQUISAS EM SAÚDE: A ESCOLHA DO TESTE ESTATÍSTICO
QUALITY EVALUATION CRITERIA FOR HEALTH RESEARCH RESULTS: STATISTICAL TESTS CHOOSING

Reinaldo José Gianini*

RESUMO

O objetivo desta atualização é aperfeiçoar o poder da crítica às publicações que propõem novas condutas, fornecendo alguns critérios para a verificação do acerto na escolha dos testes estatísticos. Nesse sentido, é focalizada na qualidade estatística da evidência clínico-epidemiológica. Recorreu-se à literatura e realizou-se uma pesquisa de *softwares* indicados para análises estatísticas, tendo sido investigadas suas sugestões e recomendações. Após estas etapas, os resultados foram organizados de modo a oferecer um algoritmo para a escolha de testes estatísticos, com base na identificação das seguintes características dos estudos de saúde: tipo de variável, número de variáveis, número de amostras, tamanho da amostra, dependência entre amostras, distribuição dos dados (paramétrica ou não) e função matemática entre variáveis.

Descritores: estatística, avaliação, pesquisa em saúde, software.

ABSTRACT

This update objective is empowering critics over publications that consider new directions in health. Some criteria for the verification of the rightness in the choice of statistical tests are supplied, focusing in the clinical and epidemiological evidence quality. Literature was appealed to it and was become fulfilled an indicated research of software's for statistical analyzes, having been investigated its suggestions and recommendations. After these stages the results had been organized in order to offer an algorithm for the choice of statistical tests, identifying the following health studies characteristics: variable type, number of variables, number of samples, sample size, dependence between samples, data distribution (parametric or not), and mathematical function between variables.

Key-words: statistics, evaluation, health research, software.

INTRODUÇÃO

Vivemos na era da informação. No futuro, provavelmente, este início de século 21 será lembrado pela verdadeira explosão na quantidade de meios e variedade de informação assim como pela enorme expansão do acesso à informação em nível internacional. O que traz grandes vantagens, por exemplo, posso ler um artigo recém-publicado em outros continentes a custo zero, via Internet, entrando na United States National Library of Medicine.¹ Mas também traz problemas: nunca foi tão difícil separar informação útil, válida, fidedigna daquela informação espúria, falsa. Quase não há mais filtros à publicação; a divulgação é livre. Especialmente nós, médicos,

somos bombardeados nos congressos ou por propagandistas da indústria farmacêutica com novidades ou descobertas recentes. Por mais cuidadosos que sejamos no esforço de verificar a validade dessas novas medidas diagnósticas ou terapêuticas, freqüentemente não encontramos metanálises,² recomendações³ ou consensos⁴ que possam nos orientar. Em várias situações, o conhecimento que temos das ciências básicas (fisiologia, patologia, farmacologia, epidemiologia) não se mostra suficiente para tomarmos uma decisão sobre qual é a melhor conduta. E inúmeras vezes, o novo contradiz nossa experiência particular ou aqueles resultados que conhecemos numa determinada série de casos.

OBJETIVOS

O objetivo desta atualização é fornecer alguns critérios para a checagem da correção na escolha dos testes estatísticos mais apropriados em cada tipo de estudo para aperfeiçoar o poder da crítica às publicações que propõem novas condutas. Nesse sentido, é focalizada na qualidade estatística da evidência clínico-epidemiológica.^{5,6,7}

MÉTODOS

Para a elaboração do presente trabalho, além de recorrer-se à literatura, foi feita uma pesquisa de *softwares* para emprego na análise estatística, tendo sido investigadas suas sugestões e recomendações.^{8,9,10,11} Após estas etapas, de revisão e investigação, os resultados obtidos foram organizados de modo a oferecer um algoritmo para a escolha de testes estatísticos.

RESULTADOS E DISCUSSÃO

O quadro procura sintetizar exatamente isso: apresenta alguns dos testes estatísticos mais recomendados segundo determinadas características das pesquisas de saúde. Pode-se decidir se o teste empregado em determinado estudo foi o mais apropriado e, portanto, se seus resultados têm qualidade no que diz respeito ao tratamento estatístico a partir da identificação das características que se seguem:

1. Tipo de variável

Variável estatística é definida como “um atributo mensurável que tipicamente varia entre indivíduos”. São classificadas em dois grupos: quantitativas (ou numéricas) e qualitativas (ou categóricas).

Rev. Fac. Ciênc. Méd. Sorocaba, v. 10, n. 3, p. 1-4, 2008

* Professor do Depto. de Medicina - CCMB/PUC-SP -, pesquisador LIM-39 HC/FMUSP.

Recebido em 11/7/2008. Aceito para publicação em 14/8/2008.

Contato: rgianini@pucsp.br

As variáveis quantitativas são numericamente mensuráveis como, por exemplo, a idade, o número de filhos. Subdividem-se em: quantitativas contínuas, aquelas que assumem valores dentro de um conjunto contínuo, tipicamente os números reais (por exemplo, o peso ou a altura de uma pessoa); e quantitativas discretas aquelas que assumem valores dentro de um espaço finito ou enumerável, tipicamente números inteiros (por exemplo, número de consultas). As variáveis qualitativas não podem ser mensuráveis numericamente. Subdividem-se em: qualitativas ordinais, aquelas que podem ser colocadas em ordem, por exemplo, a classe social (A, B, C, D, ou E) ou a variável “peso” medida em três níveis (pouco pesados, pesados, muito pesados); e qualitativas nominais aquelas que não podem ser hierarquizadas ou ordenadas, como a cor dos olhos ou o local de nascimento.¹² Em situações muito especiais, o pesquisador necessita transformar uma variável categórica em numérica para viabilizar a análise, atribuindo códigos numéricos a cada categoria. E pode ser necessário fazer o contrário, transformar variáveis numéricas em categóricas, construindo classes (por exemplo: idade de 10 a 19, de 20 a 29, etc.). Nessas situações, a interpretação dos resultados deve ser extremamente cuidadosa, não se perdendo de vista a distribuição original dos dados, em especial no que tange à análise de tendências, e não se esquecendo que o número atribuído à alguma categoria ou classe não é um valor, mas apenas um código.¹³

2. Número de variáveis incluído no teste (1, 2 ou k)

As análises univariadas incluem somente 1 variável em um teste estatístico para comparar amostras. Esses estudos podem estar interessados em analisar: diferenças entre proporções, médias, variâncias ou medianas; associações; aderência; tendência ou concordância.

As análises bivariadas incluem 2 variáveis em um teste estatístico. Podem empregar além dos testes constantes nas análises univariadas para comparação entre amostras, testes para investigar a relação entre variáveis, tais como a correlação simples, a regressão simples, a interação ou modificação de efeito.

As análises multivariadas incluem k variáveis (3 ou mais) em um teste estatístico. Podem analisar, além de todos os aspectos anteriores, a correlação parcial (matrizes de correlação) e a regressão múltipla entre k variáveis.¹⁴

3. Número de amostras incluído no teste (1, 2 ou k)

Uma amostra estatística consiste de um conjunto de indivíduos. Estudos típicos de uma amostra são o de série de casos e os estudos transversais. Para validar seus resultados é necessária a comparação com dados descritos na literatura geralmente assumidos como parâmetro populacional.¹²

O teste de Kolmorov-Smirnov para 1 variável constitui uma situação especial, pois a partir de 4 categorias pode-se testar se suas frequências respeitam ordem crescente ou decrescente sem comparação com o parâmetro.⁹

Duas amostras ocorrem, tipicamente, tanto nos estudos observacionais (caso-controle, coorte) quanto experimentais (ensaios clínicos). Nos estudos de caso-controle, uma amostra é o grupo de estudo que apresenta o desfecho (doença ou condição de interesse) representado pelos casos, e outra amostra é o grupo de comparação, necessário à validação dos resultados, que são indivíduos que não apresentam o desfecho (doença ou condição de interesse) representado pelos controles. Nos estudos de coorte, analogamente, temos o grupo de

estudo que está exposto ao fator de risco (ou de proteção) e o grupo controle que é representado por indivíduos não expostos ao mesmo fator. Nos ensaios clínicos, a exposição a determinado fator é experimental, ou seja, introduzida pelo pesquisador, que nos estudos de melhor qualidade os distribui aleatoriamente.¹⁵

Os estudos ecológicos e os ensaios na comunidade incluem vários agregados populacionais (geográficos ou temporais), comumente apresentando k amostras.¹⁶

4. Dependência ou independência entre amostras

O termo dependência (ou dependentes) na literatura relacionada à estatística aparece com diferentes conceitos. No presente trabalho utilizamos o termo para denotar medidas (ou observações) repetidas para uma mesma variável, nos mesmos indivíduos, realizadas em diferentes tempos, diferentes situações ou por diferentes observadores. Um exemplo típico são os ensaios clínicos com medidas antes e depois da intervenção (duas amostras dependentes). Mas também existem estudos que apresentam séries de medidas repetidas, aqui registrados como k amostras dependentes.¹⁷

5. Tamanho da amostra (n = 30; n > 30)

Geralmente, as variáveis quantitativas contínuas assumem distribuição normal (Gaussiana) quando $n > 30$ (maiores frequências próximas à média, 68,5% da frequência dentro do intervalo média ± 1 desvio-padrão).

Em amostras menores seguem a distribuição t de Student (distribuição de probabilidade para pequenas amostras). As variáveis quantitativas discretas podem assumir as distribuições anteriores (ditas paramétricas) ou não. As variáveis categóricas geralmente seguem a distribuição qui-quadrado para amostras de $n > 30$ e a distribuição Binomial (no caso das variáveis dicotômicas, tipo sim ou não) ou de Fisher (nas pequenas amostras).¹⁷

6. Distribuição dos dados (paramétrica ou não)

Diz-se que a distribuição é não-paramétrica quando os valores são muito discrepantes entre si, de modo que a média (e o desvio-padrão) não representa a tendência central e dispersão da amostra (nesses casos melhor representados pela mediana e percentis). São denominadas distribuições livres. O mais correto é testar-se a distribuição dos dados antes de se decidir pelo teste de hipótese paramétrico ou não (por exemplo, Shapiro-Wilk). Nas distribuições não-paramétricas, os valores originais dos dados são transformados em *rankings* (1 para o menor valor, 2 para o seguinte, n para o maior) e os testes estatísticos são aplicados aos valores ranqueados, e não aos valores originais. Os testes não-paramétricos são menos precisos.¹⁸

7. Função matemática entre as variáveis

Nem sempre $f(x)$ é uma reta onde $y = a + bx$. Em outras palavras, nem sempre a relação entre duas variáveis ocorre de modo linear. Nas variáveis dicotômicas é bem conhecida a função exponencial, tratada, portanto, não como regressão linear, mas como regressão logística. Em toda regressão, simples ou múltipla, devemos checar qual é a melhor função matemática que se aplica ao problema (linear, exponencial, logarítmica, polinomial, quadrática, etc.). Os melhores estudos que tratam de regressão devem fazer constar na metodologia este procedimento, denominado Ajuste de Curvas.⁹

Identificadas essas características da pesquisa, se o teste utilizado não se encontra entre os recomendados, procure pela justificativa. Pode ser que esta situação não anule completamente

o valor do estudo, apenas enfraqueça sua qualidade. De qualquer modo, este é o sinal para ser prudente: procurar ou aguardar novas evidências antes de mudar de conduta.

Quadro. Alguns testes estatísticos mais recomendados segundo determinadas características da pesquisa em Saúde

NÚMERO DE AMOSTRAS	TIPOS DE VARIÁVEL / DADO*	NÚMERO DE VARIÁVEIS ANALISADAS NO TESTE		
		1 variável	2 variáveis	3 variáveis
1 amostra	Nominais	Binominal, Fischer, Qui-quadrado Aderência	Fisher, Qui-quadrado	Qui-quadrado de Mantel-Haentzel, Regressão Logística
	Ordinais	Kolmorov-Smirnov #, Qui-quadrado Aderência, Fischer	Correlação de Kendall Qui-quadrado Tendência linear	Qui-quadrado de Mantel-Haentzel, Regressão Logística
	Numéricos	Teste Z (n>30), Teste t (n≤30)	Correlação de Pearson, regressão simples (linear ou não)	Regressão múltipla (linear ou não), Componentes principais (CP)
	Não-paramétricos	Mann-Whitney §, Teste das medianas	Spearman, Regressão não-paramétrica (linear ou não)	Regressão múltipla não-paramétrica (linear ou não), CP
2 amostras	DEPENDENTES	Nominais	McNemar	Modelos Mistos de Regressão (linear ou não) ajustados para a dependência, p.e., Cronbach
		Ordinais	Kappa	
		Numéricos	Teste t pareado	
		Não-paramétricos	Sinais, Wilcoxon	
	INDEPENDENTES	Nominais	Fischer, Qui-quadrado	Qui-quadrado de Mantel-Haentzel, Regressão Logística, análise de interação ou modificação de efeito
		Ordinais	Kolmorov-Smirnov, Qui-quadrado para tendência Fisher	
		Numéricos	Teste t, teste Z, ANOVA (1 critério)	Regressão múltipla (linear ou não)
		Não-paramétricos	Mann-Whitney, Teste das medianas	Regressão múltipla não-paramétrica (linear ou não), p.e., LAV
K-amostras	DEPENDENTES	Nominais	Cochran	Modelos Mistos de Regressão (linear ou não) ajustados para a dependência, p.e., Cronbach
		Ordinais	Cochran	
		Numéricos	ANOVA (2 critérios)	
		Não-paramétricos	Friedman	
	INDEPENDENTES	Nominais	Qui-quadrado	Qui-quadrado de Mantel-Haentzel, Regressão Logística
		Ordinais	Qui-quadrado, tendência	Correlação de Kendall
		Numéricos	ANOVA (1 critério)	Regressão múltipla (linear ou não)
		Não-paramétricos	Kruskal-Wallis	Regressão múltipla não-paramétrica (linear ou não)

* Dados quantitativos classificados como categóricos nominais (dicotômicos ou não) ou ordinais; dados quantitativos classificados como numéricos (paramétricos) ou não-paramétricos # auto-referido § é necessário reconhecer a distribuição completa de dados da população referência à qual se quer comparar a amostra.

CONCLUSÕES

Em conclusão, apesar de a Estatística ser um “mar de testes”, mesmo não sendo um especialista na área pode-se navegar com relativa segurança se compreendermos algumas diretrizes básicas. Procuramos dar início a este processo. Mas é claro que o assunto não se esgota aqui. É recomendável a leitura sobre os testes citados, cuja maioria encontra-se descrita com exemplos nos livros básicos dirigidos a iniciantes na área.^{12,17,18}

REFERÊNCIAS

1. Pubmed [base de dados na Internet]. Washington (DC): US National Library of Medicine, National Institutes of Health. c1993 [acesso em 15 maio 2008]. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/>
2. Cochrane BVS [base de dados na Internet]. São Paulo: Bireme [acesso em 15 maio 2008]. Disponível em: <http://cochrane.bvsalud.org/portal/php/index.php>
3. Associação Médica Brasileira [homepage na Internet]. São Paulo: AMB; 2000 [acesso em 15 maio 2008]. Projeto Diretrizes. Disponível em: <http://www.projetoDiretrizes.org.br/>
4. Agency for Healthcare Research and Quality [homepage na Internet]. Rockville (MD): AHRQ; [acesso em 15 maio 2008]. U.S. Preventive Services Task Force (USPSTF). Disponível em: <http://www.ahrq.gov/clinic/>
5. Strobe Statement Strengthening the Reporting of Observational studies in Epidemiology [homepage na Internet] [acesso em 15 maio 2008]. Disponível em: <http://www.strobe-statement.org/index.html>
6. Rouquayrol MZ, Almeida Filho N. Introdução à epidemiologia. 4ª ed. Rio de Janeiro: Guanabara-Koogan, 2006.
7. Consort Group. Consort Statement [homepage na Internet] [atualizada em 22 oct 2007; acesso em 15 maio 2008]. Disponível em: <http://www.consort-statement.org/index.aspx?o=1011>
8. Stata: Data Analysis and Statistical Software [base de dados na Internet]. Texas (US): StataCorp. c1996-2008 [acesso em 15 maio 2008]. Disponível em: <http://www.stata.com/>
9. Ayres M, Ayres Junior M, Ayres DL, Santos AS. BioEstat 4.0: aplicações estatísticas nas áreas das ciências biomédicas. Brasília: Ministério da Ciência e Tecnologia; 2005.
10. Centers for Diseases Control and Prevention. Epi Info [base de dados na Internet]. Atlanta: CDC [atualizada em 19 ago 2008; acesso em 22 ago 2008]. Disponível em: www.cdc.gov/epiinfo/
11. SPSS Statistics Base 17.0 SPSS - Statistical Package for the Social Sciences [base de dados na Internet]. Chicago: SPSS; c2008 [acesso em 22 ago 2008]. Disponível: <http://www.spss.com.br/>
12. Vieira SM. Introdução à bioestatística. 3ª ed. Rio de Janeiro: Campus; 1998.
13. Spiegel MR. Estatística. 3ª ed. São Paulo: Makron; 1993.
14. Rothman KJ, Greenland S. Modern epidemiology. Philadelphia: Lippincott-Williams & Wilkins; 1998.
15. McMahon B, Pugh TF. Epidemiology: principles and methods. Boston: Little Brow; 1970.
16. Beaglehole R, Bonita R, Kjellstion T. Epidemiologia básica. Rio de Janeiro: Santos; 1996.
17. Motta VT; Wagner MB. Bioestatística. 2ª ed. Caxias do Sul: EDUCS; 2006.
18. Triola MF. Introdução à estatística. 9ª ed. Rio de Janeiro: LTC; 2008.



REVISTA DA FACULDADE DE CIÊNCIAS MÉDICAS DE SOROCABA

Agradecemos a colaboração da Associação dos Docentes do CCMB/PUC-SP

Diretoria

Enio Marcio Maia Guerra
João Luiz Garcia Duarte
Celeste Gomes Sardinha Oshiro
José Eduardo Martinez
Dirce Setsuko Tacahashi
Nelson Boccato Jr.