



Artigo especial

Dimensionamento de amostras e o mito dos números mágicos: ponto de vista



C.J. Brito^a, M.E. da Silva Grigoletto^b, O. de Toledo Nóbrega^c e C. Córdova^{d,e,*}

^a Departamento de Educação Física, Centro de Ciências Biológicas e da Saúde, Universidade Federal de Juiz de Fora, Governador Valadares, MG, Brasil

^b Departamento de Educação Física, Centro de Ciências Biológicas e da Saúde, Universidade Federal de Sergipe, Aracaju, SE, Brasil

^c Programa de Pós-Graduação em Ciências Médicas, Faculdade de Medicina, Universidade de Brasília, Brasília, Brasil

^d Programa de Pós-graduação em Educação Física, Universidade Católica de Brasília, Brasília, Brasil

^e Programa de Pós-graduação em Gerontologia, Universidade Católica de Brasília, Brasília, Brasil

INFORMAÇÃO SOBRE O ARTIGO

Historial do artigo:

Recebido a 30 de junho de 2014

Aceite a 12 de fevereiro de 2015

Palavras-chave:

Tamanho de amostra
Viés

Projetos de pesquisa

Coleta de dados

Atividade física e saúde

R E S U M O

O presente texto dissertará sobre aspectos que devem ser cuidadosamente ponderados por pesquisadores com o propósito de se estabelecer, *a priori*, o tamanho amostral em estudos na área da saúde, aptidão física e desempenho humano. Entretanto, é importante salientar que o cálculo de amostras é uma ciência que exige seguras suposições realistas, anterior à sua estimativa. Por fim, observamos que, apesar dos grandes avanços tecnológicos que têm contribuído para a maior velocidade de produção científica, é de vital importância que o pesquisador tenha o compromisso em respeitar premissas que habilitam a qualidade do trabalho científico. Nesse sentido, a apropriada especificação do tamanho amostral confere validade interna ao estudo e deve ser requisito obrigatório para a aprovação de projetos de pesquisa.

© 2015 Consejería de Turismo y Deporte de la Junta de Andalucía. Publicado por Elsevier España, S.L.U. Este é um artigo Open Access sob a licença de CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Tamaño muestral y el mito de los números mágicos: punto de vista

R E S U M E N

El presente texto versará sobre aspectos que deberían ser cuidadosamente ponderados por parte de los investigadores con el objetivo de establecer, *a priori*, el tamaño muestral de estudios en áreas de la salud, actividad física y rendimiento humano. Hay que resaltar que el cálculo del tamaño de la muestra no es una ciencia exacta y, por tanto, requiere suposiciones realistas previas a su estimación. Por último, observamos que, a pesar de los enormes avances biotecnológicos que han contribuido a una más rápida producción científica, es de vital importancia que los investigadores se comprometan a respetar premisas que permitan la calidad del trabajo científico. En este sentido, una adecuada especificación del tamaño muestral confiere validez interna al estudio y deben ser un requisito obligatorio para la aprobación de proyectos de investigación.

© 2015 Consejería de Turismo y Deporte de la Junta de Andalucía. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Palabras clave:

Tamaño de la muestra

Sesgo

Proyectos de investigación

Recolección de datos

Actividad física y salud

* Autor para correspondência.

Correio eletrónico: claudioub@yahoo.com.br (C. Córdova).

A B S T R A C T

The present text will deal with aspects that should be carefully taken into account by researchers with the purpose of establishing sample size in studies in the areas of health, physical capacities and human development. It is important to highlight that sample calculation is not an exact science, and, as such, it requires realistic suppositions before estimation. It is observed that, despite great biotechnological advances that have contributed to faster scientific production, it is of vital importance that researchers commit to respect premises that enable the quality of the scientific work. In this regard, appropriate specifications of sample size grant internal validity to the study and must be a mandatory requisite for research projects approval.

© 2015 Consejería de Turismo y Deporte de la Junta de Andalucía. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introdução

Pesquisas são conduzidas para diferentes propósitos. Talvez o interesse do pesquisador seja em evidenciar diferença entre tratamentos em termos de parâmetros previamente definidos, como os efeitos adversos ou fatores de risco. Da mesma forma, estudos podem ser delineados com o propósito de investigar a associação entre características de interesse ou, em termos de estimativas populacionais, a prevalência de um evento¹. Nesse cenário, técnicas para o dimensionamento de amostras destacam-se como estratégias críticas tanto para a apropriada estimação dos parâmetros populacionais, quanto para a decisão de rejeitar ou não a hipótese nula (H_0). Resultados obtidos a partir de amostras subdimensionadas concorrem para conclusões equivocadas quando a não rejeição da H_0 está associada a um erro tipo II (sua não rejeição, embora H_0 seja falsa). Por outro lado, quando a seleção de participantes extrapola o quantitativo racional, verifica-se desperdício desnecessário de recursos e conduta ética inadequada. Ademais, eleva-se a probabilidade de se observar diferenças estatisticamente significativas, ainda que tal diferença não tenha nenhuma relevância biológica ou clínica. O presente artigo apresentará posicionamento crítico sobre a importância do cálculo amostral para conferir validade interna aos estudos na pesquisa em educação física e áreas afins.

Ponto de vista

O paradigma das ciências da saúde é constituído por teorias, métodos e técnicas de investigação, onde as ciências exatas ocupam posição de destaque, com particular ênfase a estatística aplicada². Com o advento do computador e o contínuo desenvolvimento de modernas tecnologias de *hardwares*, número crescente de poderosas ferramentas (*softwares*) tem contribuído para a maior agilidade de processamento, poder para análise dos dados e elaboração de modelos estatísticos que, potencialmente, ajudam a explicar a complexidade dos fenômenos investigados. Por outro lado, mesmo com todo o avanço científico e velocidade na divulgação desse conhecimento, ainda preservamos em nossa cultura antigos hábitos deterministas de nossos ancestrais – a predisposição/condicionamento de estabelecer relações de causa e efeito – e, desse modo, ignoramos as dimensões das incertezas intrínsecas aos métodos quantitativos². O dimensionamento de amostras é mais um exemplo dessa contradição. É muito comum em publicações científicas verificarmos que os resultados foram baseados arbitrariamente em tamanhos amostrais de conveniência, ou seja, sem qualquer fundamentação probabilística. Quando indagamos sobre a fundamentação teórica, é corriqueiro afirmar-se que 30 participantes ($n=30$) é um bom número, ou ainda, 10 indivíduos por grupo são suficientes ($n=10$), porque outro trabalho com o mesmo delineamento também utilizou esses números (i.e. *viés de confirmação*). A falha em não fundamentar o tamanho

amostral não é fato recente. Por exemplo, Moher et al.³ verificaram que em amostra constituída por 102 estudos que não rejeitaram H_0 , somente 36% apresentaram poder suficiente ($1 - \beta = 80\%$) para detectar uma relativa diferença de 50% entre grupos. Portanto, esses resultados confirmam que a apropriada especificação do tamanho amostral, particularmente na etapa de planejamento do poder do estudo (i.e. *a priori power analyses*) é comumente negligenciada. Em análises de poder *a priori*, o tamanho da amostra é calculado como função relativa do poder exigido ($1 - \beta$), o nível de significância pré-especificado (α) (fig. 1A) e a magnitude mínima do efeito populacional (i.e. *effect size*) a ser detectado com probabilidade ($1 - \beta$)⁴. Quando a magnitude do efeito cresce, a relação entre as probabilidades para os 2 tipos de erros também é modificada (fig. 1B).

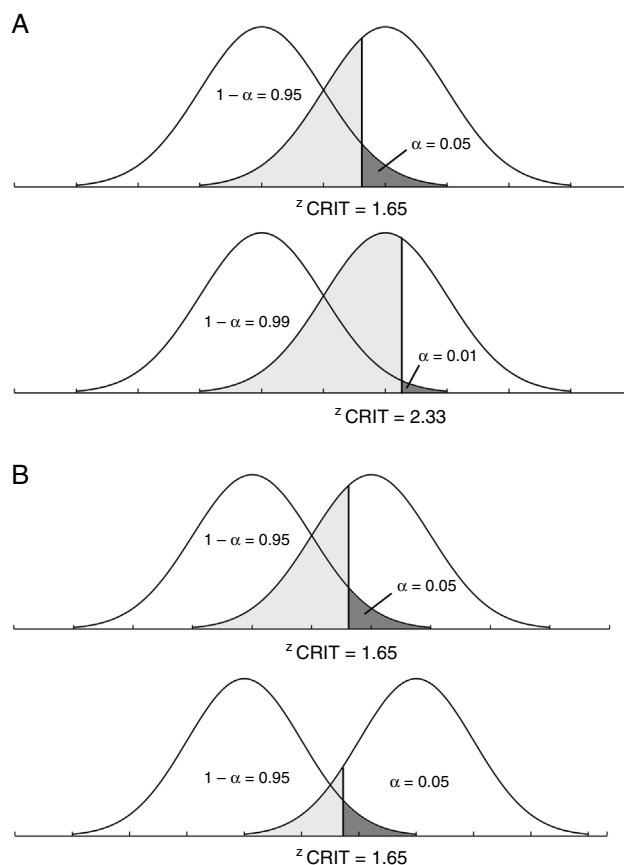


Figura 1. A. Representação gráfica ilustrando a relação entre valores de (α) e (β). B. Representação gráfica que ilustra a relação entre as probabilidades para os erros tipo I e tipo II quando a magnitude dos efeitos (*effect size*) é alterada.

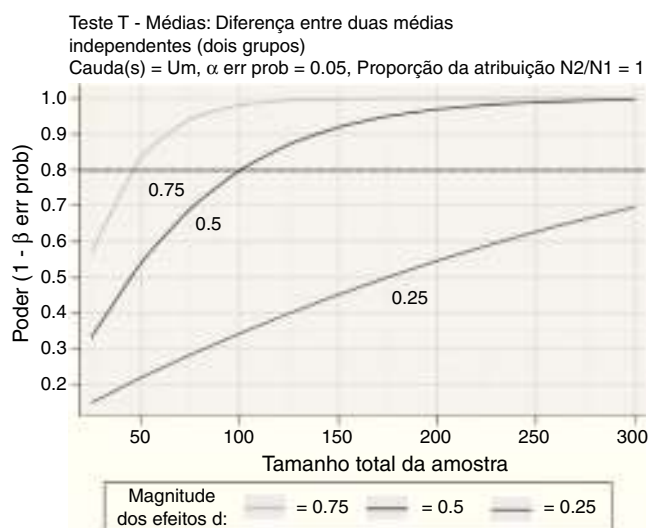


Figura 2. Exemplo da influência da magnitude do efeito (d) sobre o tamanho da amostra e do poder do teste ($1 - \beta$), $\alpha = 0.05$. Utilizou-se o software G*Power versão 3.1.9.2 para um poder de 80% é necessária uma grande amostra ($n > 300$) para se detectar um pequeno efeito ($d = 0.25$) em relação a efeitos de médio ($d = 0.50$) a grande magnitude ($d = 0.75$) ($n = 100$ e $n < 50$, respectivamente).

É importante ressaltar que existem cenários onde o dimensionamento de amostras *a priori* não é algo factível. Por exemplo, em estudos que contemplam amostras muito seletas e, portanto, pequenas, como no caso de doenças raras ou com atletas de alto rendimento. Nesse panorama, os cálculos amostrais *a posteriori* (i.e. *post hoc power analyses*) são justificados e pertinentes, uma vez que disponibilizam para a literatura especializada estimativas de parâmetros – médias, medidas de dispersão e precisão adequada para a magnitude mínima do efeito – para futuros trabalhos. Para grande parte das medidas contínuas, convencionou-se o limiar de 0.20 para a diferença de médias e de 0.10 para correlações⁴.

Da mesma forma, é fundamental que pesquisadores que trabalham, tanto com modelos humanos, quanto animais, justifiquem o número de indivíduos a serem incluídos ao Comitê de Ética em Pesquisa, entidade responsável pelo parecer da proposta da pesquisa, cuja responsabilidade inclui a análise das justificativas técnicas sobre a casuística sugerida ao estudo⁵. Nesse sentido, como grande parte dos trabalhos que contemplam desenhos experimentais ou observacionais envolve a formalização de testes de hipóteses, dados coletados a partir de experimentos exploratórios ou de estudo piloto devem ser utilizados a fim de responder a algumas poucas questões, porém, necessárias para a apropriada estimativa do tamanho amostral:

- Qual a margem de erro, absoluta ou relativa, desejada na variação dos resultados?
- Qual a probabilidade ou o risco associado com os resultados desejados?
- Qual a estimativa da dispersão das variáveis envolvidas na investigação?

Embora a análise de poder do teste seja indispensável para decisões estatísticas fundamentadas, somente a partir da década de 80 é que gráficos⁶ e tabelas de poder⁴ foram suplementados por programas de análises de poder de testes. Alguns disponibilizados gratuitamente, como o programa G*Power (fig. 2), e com evidentes vantagens aos usuários: eficiência, precisão e facilidade de operacionalização⁷. Alternativa mais recente e de livre acesso é a planilha de Excel[®] para cálculo amostral para inferências baseadas em magnitudes⁸.

Nesse cenário, é notório o grande avanço que essa tecnologia tem proporcionado para a aplicação prática desse conhecimento. Entretanto, é importante ressaltar que a agilidade e a aparente facilidade na seleção dos itens exibidos em caixas de textos (programas que rodam em plataforma Windows) dependem de prévio *background* teórico baseado em algoritmos numéricos e métodos estatísticos. Ademais, devemos ter em mente a responsabilidade e o compromisso em respeitar premissas que habilitam a qualidade do trabalho científico, não somente para a publicação em periódicos e comunicados científicos, como também com exigência ética.

Recomendações finais

Em síntese, a apropriada especificação do tamanho amostral confere validade interna ao estudo e deve ser requisito para a aprovação de protocolos de pesquisa e a publicação de trabalhos em revistas científicas. Deve ser realizada no momento do planejamento do estudo por cientistas com domínio do tema. Em trabalhos publicados, os autores devem fornecer informações detalhadas quanto aos cálculos e procedimentos utilizados, a fim de justificar a casuística do estudo.

Conflito de interesses

Os autores declaram não haver conflito de interesses.

Agradecimentos

Daniel A. Boulosa pelas contribuições na leitura do texto.

Bibliografia

- Nayak BK. Understanding the relevance of sample size calculation. *Indian J Ophthalmol.* 2010;58:469-70.
- Pereira JC. Tamanho de amostra: uma necessidade real ou um capricho cultural. *Arteriola.* 2002;4:13-6.
- Moher D, Dullberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. *JAMA.* 1994;272:122-4.
- Cohen J. *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ, England: Lawrence Erlbaum Associates; 1988. p. 590.
- Dell RB, Holleran S, Ramakrishnan R. Sample size determination. *ILAR J.* 2002;43:207-13.
- Scheffe H. *The analysis of variance.* New York: John Wiley & Sons; 1999.
- Goldstein R. Power and sample size via MS/PC-DOS computers. *Am Estetician.* 1989;43:253-60.
- Hopkins W. Estimating sample size for magnitude-based inferences. *Sportscience.* 2006;10:63-70 [12/05/2014]. Em [<http://www.sportsci.org/2006/wghss.htm>].