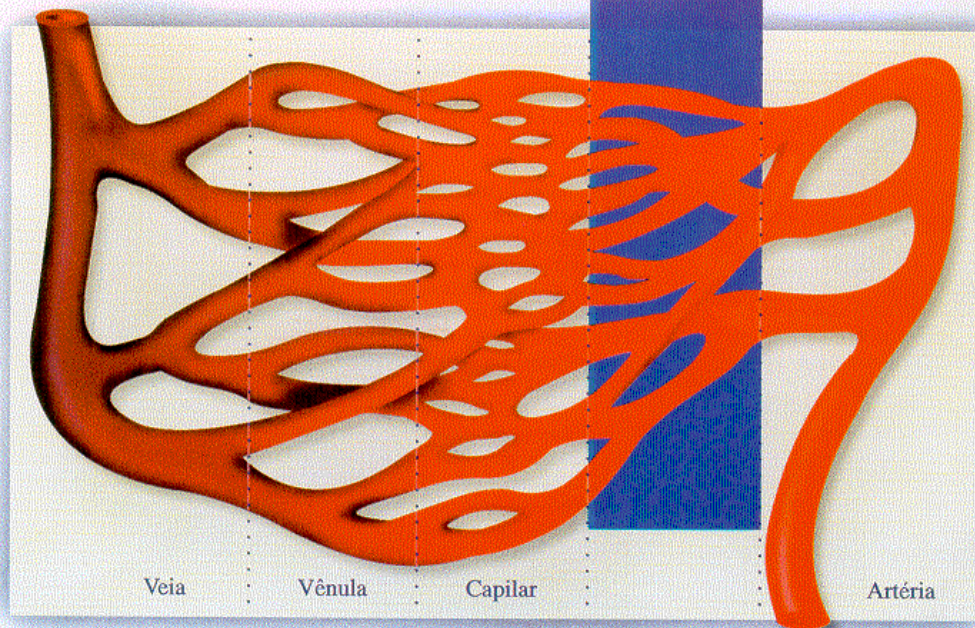


Arteriola 2002, 4(1):13-16

Arteriola

ÓRGÃO DE DIVULGAÇÃO CIENTÍFICA DA SEÇÃO DE HIPERTENSÃO
ARTERIAL E NEFROLOGIA DO INSTITUTO "DANTE PAZZANESE"
DE CARDIOLOGIA DO ESTADO DE SÃO PAULO
E DA FUNDAÇÃO ADIB JATENE

Volume 4 - nº1 - 2002



Tamanho de amostra: uma necessidade real ou um capricho cultural?

Júlio C. R. Pereira*

Nos anos 50 do século XX, um físico – Thomas Kuhn – trabalhando em sua pós-graduação em Física Teórica, veio a transformar-se num maiores filósofos da ciência de nossos tempos. Argüindo-se sobre a teoria do conhecimento científico, Kuhn desenvolveu o conceito de paradigma científico¹, que é o conjunto de valores de natureza metafísica, epistemológica e ética que, historicamente dados, condicionam os meios e formas de pensar, investigar e conhecer da ciência. Segundo Kuhn, o conhecimento científico não é cumulativo, ou seja, o novo não acrescenta ao antigo, mas o novo substitui o antigo num processo que ele chamou de revolução científica. Quando historicamente esgotam-se as potencialidades de um dado paradigma, um novo surge para proporcionar uma visão totalmente nova da natureza. Este movimento revolucionário não implica, no entanto, em negação do conhecimento antigo – apenas restringe-o a seus limites paradigmáticos. Exemplo de uma revolução científica é o nascimento da mecânica quântica quando do esgotamento das possibilidades da estrutura paradigmática da mecânica clássica.

O paradigma da ciência médica atual inclui um referencial epistemológico – de teorias, métodos, técnicas, procedimentos – onde as matemáticas se destacam, com particular ênfase para a estatística e a probabilidade. Quando se fala em *medicina baseada em evidências*, fala-se em evidências com lastro neste referencial epistemológico: quantificação de relações de causa e efeito. Este ideário decorre da reificação do conhecimento de base empírica e seu suposto tratamento objetivo, em detrimento do conhecimento de natureza puramente racional.

Estes dois modelos alternativos de produção do conhecimento, são designados por Crosby² como o modelo venerável (razão pura) e novo modelo (quantificação e experimentação empírica). Até meados do século XIII a ciência não se preocupava com a quantificação até porque entendia-se que o empírico não era necessariamente o real, mas apenas sua percepção através dos sentidos humanos. Segundo Crosby, é a partir desta época que o interesse em medidas começa a crescer na humanidade, afetando todo tipo de atividade humana, não só as ciências – a tudo se busca quantificar: tempo, espaço, música, pintura, comércio, segundo os exemplos que explora em diferentes capítulos de seu livro. O novo modelo estabelece-se plenamente na renascença, constituindo-se em marco histórico os trabalhos de Galileu Galilei, particularmente seu “*Dialogo sopra I massimi sistemi*” (Diálogo sobre os sistemas máximos) que, segundo seus biógrafos³, ele queria ter chamado “*Discorsi e dimostrazioni matematiche intorno a due nuove scienze, attinenti alla meccanica e i movimenti locali*” (Discursos e demonstrações matemáticas em torno de duas novas ciências atinentes à mecânica e aos movimentos locais), título que deixa claro que se trata de um modelo novo apoiado em demonstrações matemáticas. Um novo paradigma se

* Professor do Depto. de Epidemiologia da Fac. de Saúde Pública da USP

estabelecia para as ciências instrumentalizando uma notável revolução do conhecimento que constitui a base das ciências modernas, embora Galileu não entendesse que suas *novas ciências* se opusessem ao modelo venerável: - “ Sou peripatético!”, insistia ele.

Dentro deste paradigma, a estatística se estabeleceu inicialmente com técnica de registro de dados, com função puramente atuarial. No entanto, os exercícios descritivos da estatística começaram a sugerir padrões de comportamento de fenômenos e, progressivamente, ela passou a ser instrumento para entendimento e explicação dos fenômenos que descrevia. No século XVIII, a burocracia estatal passa a regularmente coletar dados estatísticos da população para fins de taxaço, recrutamento militar e dimensionamento do Estado. Ian Hacking⁴ conta que por volta de 1.685, o grande matemático Leibniz já recomendava a criação de um escritório central de estatística para o Estado Prussiano, enquanto algo semelhante também ocorria na Inglaterra. Mais tarde, em 1.880, as estatísticas sobre migração e presença de judeus no Novo Império Germânico vão gerar grandes disputas de interpretação num ambiente de grande anti-semitismo, que por sua vez também, embora fenômeno social, encontra ai registros históricos em informações estatísticas.

O modelo empírico e experimental do novo paradigma dirige as idéias iluministas da renascença para uma interpretação determinista da natureza, a qual pretende que providos os dados sobre os fenômenos universais e ajustado um adequado modelo de processamento e análise destes dados, a natureza estaria perfeitamente determinada e seu conhecimento completo. Laplace dizia que se lhe fosse provida uma inteligência capaz de armazenar todas as informações do universo, tudo seria previsível em tempo e espaço, desde o movimento dos astros até a estrutura do átomo. Este determinismo que permeava o ideário do culto gerava um ambiente que tornava vulgar qualquer idéia de chance, de aleatoriedade, de incerteza. No entanto, à medida que a análise de dados estatísticos progride para o reconhecimento de padrões de distribuição – *a curva de distribuição de observações com a forma de sino descrita por Gauss vai ser chamada de Normal pelos ingleses, sugerindo que o normal deixa de ser o correto para ser o ordinário, o medíocre, o ajustado a uma tendência* – a incerteza vai se revelando empiricamente e se impondo à consideração dos cientistas: o próprio Laplace vai acabar dando importantes contribuições ao desenvolvimento das teorias de probabilidade, anotando que probabilidade seria em parte resultado de nosso conhecimento e em parte resultado de nossa ignorância.

A medicina do século XIX, embora alicerçada em quantificação detalhada de experimentos, desprezava qualquer idéia de tendência estatística ou chance probabilística. Claude Bernard⁵ ironizava o raciocínio estatístico referindo-se a ele como algo semelhante a contar quantas pessoas entram numa casa e quanta fumaça sai da chaminé. Mas também, ao mesmo tempo cresciam as tentativas de aplicação do método estatístico para produção do conhecimento médico: em 1.836, Jean Civiale⁶ ganhou o prêmio Montyon na França com um trabalho que comparava duas técnicas cirúrgicas para o tratamento de litíase biliar, demonstrando a superioridade de um

método sobre o outro com base no argumento de que a razão de óbito era de 1.024 : 5.443 num método e de 7 : 307 noutro método. O júri era presidido pelo então já notável matemático e estatístico S. D. Poisson e o veredicto do prêmio gerou tantas controvérsias que, embora o prêmio tenha sido dado a Civiale, recomendava-se que abordagens com cálculo de probabilidades fossem evitadas para estudos médicos.

As ciências da saúde de nossos dias incorporaram as técnicas estatísticas e probabilísticas como estratégia metodológica habitual na investigação científica mas, no entanto, não abandonou seu ideário determinista e tende a ignorar a dimensão de incerteza que é intrínseca a este método. A discussão em torno do tamanho da amostra, que assombra os pesquisadores médicos, é apenas um exemplo particular desta contradição: de técnicas estatísticas peneiras de incerteza busca-se a certeza de um número que se devidamente respeitado supõe-se – erradamente – uma verdade definitiva seja alcançada.

A preocupação da estatística no que tange a amostras que pretendem representar um dado universo diz respeito a sua representatividade, sua isenção de tendenciosidades, de erros sistemáticos. A ela interessa principalmente o que se chama de **Plano Amostral**: condições de entrada ou exclusão de dados, estratégias de estratificação da coleta de dados, definições de estágios de coleta de dados. São aspectos aos quais os estatísticos dão grande elaboração teórica, com muitos deles dedicando-se à amostragem como especialidade dentro da estatística⁷. No entanto, o plano amostral diz respeito às características do objeto estudado e seu conhecimento é ônus do pesquisador médico – raramente um estatístico conhecerá variáveis que estratificam o evento que ele esteja estudando, *vg.* sexo e tumores gonadais, ou sua distribuição geográfica, para sugerir estratos e estágios amostrais. Esta preocupação, que é atribuição do próprio do pesquisador médico, no entanto, é geralmente preterida em favor do cálculo de um número de tamanho de amostra que ele espera seja o aval definitivo para seu trabalho.

A preocupação da estatística está em dimensionar a incerteza da informação que gera. Por isso, calculada uma estimativa de medida, ela busca dimensionar o erro desta medida. Nesta avaliação, informações sobre a variabilidade das medidas originais e número de observações realizadas são itens levados em conta para estimativa do erro e uma fórmula matemática é então concebida onde o erro passa a ser função destas informações. Dada esta fórmula matemática, operações algébricas podem sempre inverter os papéis sobre o que é informação e o que é estimado e, desta forma, o número de observações pode ser estimado se o erro for informado. No cálculo de um tamanho de amostra, o pesquisador informa o erro com base em experiências da literatura – onde estimativas de medidas com as devidas avaliações de erro já tenham sido registradas – ou arbitra o que seja um erro aceitável para seu estudo e estima o número de observações correspondente: faz o caminho inverso ao do cálculo de erro.

A arbitrariedade do pesquisador estará sempre presente porque mesmo quando ele busque lastro na literatura, esta lhe oferecerá uma variedade de alternativas que conseqüentemente resultará em grande variabilidade do tamanho de amostra calculado. Daí que, a principal finalidade do cálculo de tamanho de amostra não é nada que se refira à confiabilidade dos resultados, mas sim ao planejamento do trabalho de pesquisa: o pesquisador assumindo determinadas premissas, quer de seu juízo próprio quer de informações de literatura, pode antever a envergadura do trabalho a que se propõe e tomar decisão sobre sua exequibilidade. Cálculo de tamanho de amostra é um instrumento de planejamento, desenho de plano amostral é um instrumento de confiabilidade e este tem **força** no conhecimento que o pesquisador tem de seu objeto e não em qualquer conhecimento que a estatística tenha sobre técnicas amostrais – estas só encontram sentido se alimentadas pelo conhecimento do pesquisador. Força aqui não é um termo ingênuo, mas diz respeito à verdade da inferência que se pretenderá fazer com o estudo realizado: em termos de lógica⁸, o método científico é indutivo (ou hipotético-dedutivo) e nestes termos não se aplicam conclusões sobre falso ou verdadeiro, mas conclusões sobre **força** da inferência (generalizações de resultados) segundo a **força** das premissas..

Para que estes comentários não resem no plano abstrato da discussão de conceitos, tomemos um exemplo. Seja a situação simples de um hipotético estudo de prevalência de Hipertensão Arterial, para o qual o pesquisador se indague sobre o tamanho da amostra. A situação original seria a da estimativa de prevalência, ou seja da **estimativa de uma proporção*** para a qual calcula-se um erro de estimativa, e a situação que se coloca é de, dado um erro de estimativa, qual seria o número de observações.

Quando estudando-se uma amostra se constata quantas pessoas apresentam uma determinada característica, tem-se uma **estimativa da proporção** de pessoas com esta característica na população da qual a amostra foi extraída. Admite-se, ou pelo menos supõe-se, que a proporção de pessoas com tal característica na população não seja exatamente a encontrada na amostra, mas algo parecido. Em outras palavras, admite-se que possa haver algum erro nesta estimativa, um erro que será diretamente proporcional às variações da presença desta característica na população (*quanto mais errática for a característica maior a chance de erro*) e inversamente proporcional ao tamanho da amostra (*quanto maior for a amostra menor a chance de erro*). Uma **medida de referência de erro** pode então ser calculada como a razão entre variabilidade – desvio padrão da proporção – e número de observações. Esta medida de referência de erro é chamada de **erro padrão de uma estimativa de proporção** e é calculado como

* A rigor uma estimativa de proporção a ser ajustada segundo características de sensibilidade e especificidade do instrumento usado para diagnóstico

$$EPP = \frac{DPP}{n} = \frac{\sqrt{np(1-p)}}{n} = \frac{\sqrt{np(1-p)}}{\sqrt{n^2}} = \sqrt{\frac{p(1-p)}{n}}$$

EPP = erro padrão da proporção
 DPP = desvio padrão da proporção
 p = proporção de casos com a característica estudada
 1 - p = proporção de casos sem a característica estudada

Habitualmente se admite como um **erro aceitável**, um erro **menor do que duas medidas de referência de erro**, ou seja dois EPP*, porque numa variação aleatória de medidas de proporções de uma amostra para outra há 95% de chances que as diferenças não superem este limite, seja para mais ou para menos. Assim, o erro é calculado como duas medidas de referência, dois EPP:

$$erro = 2EPP = 2\sqrt{\frac{p(1-p)}{n}}$$

o que significa que, **estimado** um valor para a proporção, pode-se admitir que seu **valor real** seja até um erro a mais ou a menos.

Com esta fórmula pode-se agora inverter a situação original: de uma estimativa de erro, segundo um dado número de observações e uma dada variabilidade, para uma estimativa de número de observações, segundo um dado erro e uma dada variabilidade:

$$n = 4 \frac{p(1-p)}{erro^2}$$

Se no exemplo hipotético da Hipertensão Arterial com esta fórmula o pesquisador fosse buscar apoio de informações na literatura, ele poderia encontrar informações e realizar cálculos de tamanho de amostra semelhantes aos que a tabela abaixo descreve:

Tabela 1 - Exemplos de cálculo de tamanho de amostra para um estudo de prevalência de Hipertensão arterial com base em três estudos distintos encontrados na literatura

Local	Prevalência estimada	Erro estimado	Nº calculado de observações necessárias segundo estas informações
Holanda, 2000 ⁹	13,5%	0,3%	49.845
Ludhiana, Índia, 1985 ¹⁰	14,08%	1,1%	3.841
Tanzânia, 2000 ¹¹	30,0%	4,9%	336

As divergências de *valores de n* são até 15 mil vezes! Note-se que a força das informações (*premissas*) que o pesquisador escolherá para usar (*assumir*) em seu trabalho não dizem respeito à qualidade dos estudos encontrados na literatura nem à natureza do cálculo estatístico: dizem

respeito ao conhecimento que ele tenha de seu objeto – a Hipertensão –, de quanto é pertinente ele assumir que sua população tenha características mais assemelhadas à Holanda ou à Tanzânia, de qual é a fisiopatologia da doença que condiciona diferentes taxas de ocorrências em diferentes populações.

Se o compromisso do nosso pesquisador hipotético é o de respeitar paradigmas, ele pode escusar-se do esforço de examinar dezenas de milhares de pacientes e escudar-se em premissas que o autorizam a examinar apenas algumas centenas. Ele terá cumprido o ritual que se exige dele, e seu trabalho estará habilitado para a aceitação por seus pares, sejam eles editores de revistas ou membros de um banca examinadora. Se, no entanto, seu compromisso for com o conhecimento, então estas questões, embora não lhe sejam totalmente estranhas, certamente são complementares, acessórias. Os princípios – matemáticos, estatísticos, filosóficos – que a ciência se utiliza para a produção do conhecimento científico não são o conhecimento, são instrumentos circunstanciais que têm a temporalidade dos paradigmas que os instituem.

A generalização da informação particular obtida num determinado estudo para uma regra universal não decorre do número de observações realizadas, mas da inteligência humana para aferir a força do raciocínio indutivo. Henri Poincaré, notável matemático e filósofo francês do século XIX anotava¹² que *'Princípios (vg. tamanho de amostra) são convenções e definições disfarçadas. Deduzidos de leis experimentais (da compilação de dados para a estatística) nossa mente lhes atribui um valor absoluto (valor de n !). ... não devemos nos esquecer que apenas a primeira coisa (aquela observada) é certa, todas as outras são apenas prováveis. ... mas ... é muito melhor prever com incerteza do que de todo nunca prever. Nós não sabemos, mas ainda temos que agir. Para a ação, não temos o tempo necessário para superar nossa ignorância. Precisamos, portanto, decidir sem saber. O que sei é, não que uma dada coisa seja verdadeira, mas que a melhor atitude para mim é agir como se ela fosse verdadeira'*. Tenha sido o 'valor de n' calculado ou desprezado, e ainda haverá sempre a incerteza do futuro. Mas, o compromisso com o conhecimento e com a ação dele decorrente não pode consentir a indiferença: para nosso pesquisador hipotético, intolerável seria renunciar à produção de algum conhecimento sobre a prevalência de Hipertensão e não a aceitação de um número não calculado de observações.

O compromisso do pesquisador deve ser com o conhecimento possível no seu tempo e as informações que dispõe não podem ser desprezadas em favor de um ritual, de um capricho.

* A rigor, 1,96 EPP, o que implica que na fórmula de n, onde aparece 4 a rigor deve-se ter 3,8416 (1,96²).

Para o leitor interessado em calcular tamanho de amostra, o Laboratório de Epidemiologia e Estatística (Lee) do Instituto Dante Pazzanese de Cardiologia (IDPC) mantém uma pagina na INTERNET: www.lee.dante.br/pesquisa/amostragem/amostra.html

Referências

- ¹ Kuhn TS. A estrutura das revoluções científicas. 3ª Edição. São Paulo: Editora Perspectiva; 1994.
- ² Crosby AW. The measure of reality – quantification and western society, 1250 – 1600. New York: Cambridge University Press; 1999.
- ³ Geymonat L. Galileu Galilei; tradução de Eliana Aguiar. Rio de Janeiro: Nova Fronteira; 1997.
- ⁴ Hacking I. The taming of chance. New York: Cambridge University Press; 1998.
- ⁵ Vandembroucke JP. Clinical investigation in the 20th century: the ascendancy of numerical reasoning. Lancet 1988; 352(Suppl.): 12-16.
- ⁶ Civiale J. Parallèle des divers moyens de traiter les calculeux. Paris, 1836. Apud Hacking I, op. cit.
- ⁷ Silva, N N. Amostragem probabilística : um curso introdutório. São Paulo: EDUSP; 1998
- ⁸ Kegley CW. Kegley JA. Introduction to Logic. Columbus, Ohio: Charles E. Merrill Publishing Co., 1978: 335-65.
- ⁹ Klungel OH, de Boer A, Paes AHP, Nagelkerke NJD, Seidell JC, Bakker A. Estimating the prevalence of hypertension corrected for the effect of within-person variability in blood pressure. J Clin Epidemiol 2000; 53 (11): 1158-1163.
- ¹⁰ Gupta R, al-Odat NA, Gupta VP. Hypertension epidemiology in India: meta-analysis of 50 year prevalence rates and blood pressure trends. J Hum Hypertens 1996;10(7):465-72.
- ¹¹ Edwards R, Unwin N, Mugusi F, Whiting D, Rashid S, Kissima J, Aspray TJ, Alberti KGMM. Hypertension prevalence and care in an urban and rural area of Tanzania. J Hyperten; 2000; 18 (2): 145-152.
- ¹² Poincaré H. Science and hypothesis. New York: Dover Publications Inc.; 1952: 138, 143-144, 187.