

# Análise do emprego do cálculo amostral e do erro do método em pesquisas científicas publicadas na literatura ortodôntica nacional e internacional

David Normando\*, Marco Antonio de Oliveira Almeida\*\*, Cátia Cardoso Abdo Quintão\*\*\*

## Resumo

**Introdução:** o dimensionamento adequado da amostra estudada e a análise apropriada do erro do método são passos importantes na validação dos dados obtidos em determinado estudo científico, além das questões éticas e econômicas. **Objetivo:** esta investigação tem o objetivo de avaliar, quantitativamente, com que frequência os pesquisadores da ciência ortodôntica têm empregado o cálculo amostral e a análise do erro do método em pesquisas publicadas no Brasil e nos Estados Unidos. **Métodos:** dois importantes periódicos, de acordo com a Capes (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), foram analisados, a *Revista Dental Press de Ortodontia e Ortopedia Facial (Dental Press)* e o *American Journal of Orthodontics and Dentofacial Orthopedics (AJO-DO)*. Apenas artigos publicados entre os anos de 2005 e 2008 foram analisados. **Resultados:** a maioria das pesquisas publicadas em ambas as revistas emprega alguma forma de análise do erro do método, quando essa metodologia pode ser aplicada. Porém, apenas um número muito pequeno dos artigos publicados nesses periódicos apresenta qualquer descrição de como foram dimensionadas as amostras estudadas. Essa proporção, já pequena (21,1%) na revista editada nos Estados Unidos (*AJO-DO*), é significativamente menor ( $p=0,008$ ) na revista editada no Brasil (*Dental Press*) (3,9%). **Conclusão:** os pesquisadores e o corpo editorial, de ambas as revistas, deveriam dedicar uma maior atenção ao exame dos erros inerentes à ausência de tais análises na pesquisa científica, em especial aos erros inerentes a um dimensionamento inadequado das amostras.

**Palavras-chave:** Bioestatística. Cálculo amostral. Erro do método.

**Como citar este artigo:** Normando D, Almeida MAO, Quintão CCA. Análise do emprego do cálculo amostral e do erro do método em pesquisas científicas publicadas na literatura ortodôntica nacional e internacional. *Dental Press J Orthod.* 2011 Nov-Dec;16(6):33.e1-9.

» Os autores declaram não ter interesses associativos, comerciais, de propriedade ou financeiros que representem conflito de interesse, nos produtos e companhias descritos nesse artigo.

\* Professor da Disciplina de Ortodontia da Faculdade de Odontologia da UFPA. Doutor em Ortodontia pela UERJ. Mestre em Clínica Integrada pela FOU SP. Especialista em Ortodontia pela PROFIS-USP.

\*\* Mestre e Doutor em Ortodontia pela UFRJ. Professor Titular de Ortodontia, UERJ.

\*\*\* Mestra e Doutora em Ortodontia pela UFRJ. Professor Adjunto da Disciplina de Ortodontia, UERJ.

## INTRODUÇÃO

Estudos científicos são mais confiáveis quando são cautelosamente planejados. O problema a ser investigado deve ser bem definido e operacionalizado a partir de amostras selecionadas aleatoriamente de populações apropriadas. Os métodos utilizados devem ser seguidos cuidadosamente e as medidas, obtidas através de instrumentos confiáveis e previamente testados<sup>6</sup>. Finalmente, o estudo deve ter um tamanho adequado em relação aos seus objetivos e com amostras suficientemente grandes para que um efeito considerado clinicamente importante seja, também, estatisticamente significativo.

O dimensionamento da amostra é um processo matemático de decisão, antes do início do estudo, de quantos indivíduos ou espécimes devem ser estudados<sup>5</sup>. O tamanho amostral é importante por questões éticas e econômicas<sup>7</sup>. Um estudo utilizando amostra pequena pode não produzir resultados úteis, expondo os seus participantes (amostra) a riscos desnecessários, enquanto uma amostra de tamanho excessivo usa mais fonte de recursos do que é necessário, além de expor um excessivo número de indivíduos aos riscos porventura existentes.

Infelizmente, o “distanciamento natural” dos pesquisadores e clínicos da área da saúde em relação ao entendimento da matemática e, conseqüentemente, à compreensão dos métodos estatísticos, somado à linguagem pouco acessível usada pelos estatísticos na comunicação com os profissionais da área da saúde, torna esse assunto desinteressante para a maioria dos pesquisadores clínicos<sup>12</sup>. O desconhecimento e o desinteresse aumentam a probabilidade de erros no desenho e na análise dos resultados de um estudo científico, reduzindo a confiabilidade dos dados obtidos. O estudo e a compreensão desses erros ajudariam o pesquisador a entender os motivos que levaram ao seu aparecimento e os passos necessários para tentar minimizá-los.

## CONCEITOS BÁSICOS

Pressupondo o distanciamento natural de uma considerável parte dos pesquisadores e clínicos em relação aos métodos estatísticos, será delineada uma breve revisão dos conceitos necessários para o entendimento dos temas que envolvem o cálculo para obtenção de uma amostra com tamanho adequado e dos erros inerentes ao método empregado para a mensuração dos dados coletados<sup>4,10</sup>.

### Cálculo amostral

Um pesquisador, ao realizar a análise estatística dos seus resultados, após a devida escolha de um teste estatístico, está sujeito a dois tipos de erros inerentes a essa fase. O erro tipo I (alfa) e o erro tipo II (beta)<sup>7</sup>.

Para facilitar o entendimento desses dois tipos de erros, deve-se considerar como exemplo um teste estatístico usado para definir se duas amostras apresentam diferenças significativas, e compará-lo a um julgamento que irá definir se um réu é culpado ou inocente.

### Erro tipo I ( $\alpha$ ) ou falso positivo

No final de todo um processo judicial, o júri deverá emitir o seu parecer, julgando o réu como inocente ou culpado, e levando em consideração que, a princípio, todos são inocentes até que se prove o contrário. Estatisticamente, esse princípio é definido como a “hipótese de nulidade ( $H_0$ )”, ou seja, que os dados obtidos no estudo não produziram provas suficientes para considerar as amostras como significativamente diferentes.

Ao declarar um réu como culpado, o júri poderá estar incorrendo em um erro, se, na verdade, o réu for inocente. O erro desse júri pode exemplificar, na análise estatística, o erro tipo I, ou falso positivo. No caso exemplificado, as amostras são consideradas estatisticamente diferentes, porém, na verdade, são semelhantes, ou foram obtidas da mesma população.

Na análise estatística, a probabilidade de ocorrência do erro tipo I (ou alfa) é definida pelo

valor  $p$ , obtido em qualquer teste estatístico. Portanto, quanto menor o valor  $p$ , menor será a possibilidade (probabilidade) de erro ao afirmar, no exemplo, que a diferença obtida na comparação das duas amostras é estatisticamente significativa, ou que tenha ocorrido por acaso<sup>10</sup>. É notória a grande dificuldade dos investigadores iniciantes em entender o valor  $p$ , porque não vinculam o valor obtido como sendo uma PROBABILIDADE de ERRO. Assim, quanto menor o seu valor, maior será a significância estatística analisada.

O nível de erro tipo I ( $\alpha$ ) deve ser definido antes do início do estudo e, geralmente, na Odontologia, é definido como sendo menor ou igual a 5%. Portanto, a probabilidade máxima de erro aceita para se rejeitar a hipótese de nulidade (igualdade) é de 5%, ou  $p < 0,05$ .

### Erro tipo II ( $\beta$ ) ou falso negativo

De volta ao julgamento...

Outra possibilidade de erro ocorreria se o júri declarasse o réu inocente (aceitar a hipótese de nulidade), e o réu fosse verdadeiramente culpado. Estatisticamente, a hipótese de nulidade (igualdade) seria aceita quando, na verdade, as amostras são diferentes ou foram retiradas de populações diferentes. Esse tipo de erro, conhecido como Beta ( $\beta$ ) ou falso negativo, poderá ocorrer casualmente, por probabilidade, mas é mais comum quando o tamanho da amostra é demasiadamente pequeno para alcançar uma diferença estatisticamente significativa<sup>10,13</sup>. A análise estatística não teria força suficiente (PODER) para ler uma diferença verdadeira que está ocorrendo entre as amostras analisadas<sup>6</sup>.

Comumente, depara-se, na literatura ortodôntica, com artigos científicos em que os investigadores observam uma diferença clinicamente importante entre dois grupos, porém, ao avaliar estatisticamente os dados, não conseguem obter uma diferença significativa ( $p > 0,05$ ). Assim como é muito comum a seguinte explicação: “Provavelmente, com uma amostra de maior tamanho poderia-se obter

uma diferença estatisticamente significativa entre os grupos”. Quando um investigador observa entre grupos uma diferença clinicamente importante, porém sem significância estatística, define-se o estudo como “sem poder”. Poder, em estatística, reflete a probabilidade de um estudo evitar um resultado “falso negativo”<sup>10</sup>, ou seja, no exemplo citado anteriormente, de definir as amostras como semelhantes e, de fato, elas serem semelhantes.

Assim, poder em estatística expressa a probabilidade de detectar um efeito verdadeiro. Convencionalmente, o poder tem sido estabelecido como 80% ou 90% (0,8 ou 0,9) e é igual a  $1 - \beta$ . Portanto, para o poder ser no mínimo igual a 0,8, a probabilidade de ocorrer o erro tipo II ( $\beta$ ) não deve ser maior do que 20% (0,2). Poder =  $1 - 0,2 = 0,8$ .

Amostras pequenas reduzem o poder de um estudo; contudo, amostras demasiadamente grandes proporcionam, geralmente, significância estatística entre os grupos<sup>7</sup> estudados, mesmo quando a diferença é muito pequena e considerada clinicamente sem importância.

Essa revisão dos conceitos acerca dos erros comuns em estatística reforça a necessidade de se desenharem estudos aonde o tamanho das amostras seja adequado para evitar falsas interpretações dos resultados e, conseqüentemente, tratamentos clínicos impróprios. Entretanto, além dos níveis  $\alpha$  e  $\beta$ , outros fatores fazem parte da receita que define o tamanho amostral<sup>1,6,10</sup>, sendo eles:

a) O efeito clínico mínimo que o pesquisador deseja constatar na variável principal (primária) do estudo desenvolvido. Quanto menor o efeito a ser examinado, maior deverá ser o tamanho da amostra ( $n$ ).

b) Como os dados obtidos vão ser mensurados (escala contínua, ordinal, ou nominal). Dados nominais (ex: presente x ausente/ Classe I, Classe II e Classe III) ou ordinais (ex: dor severa, moderada, suave) exigem amostras com maior tamanho, quando comparados a dados numéricos contínuos (ex: medidas cefalométricas).

c) O tipo de teste estatístico a ser usado, e se uni ou bilateral. O tipo de teste é escolhido a partir do tipo de variável mensurada e do tipo de distribuição dos dados, normal ou livre. Testes paramétricos, que utilizam dados contínuos com distribuição normal (ex: teste t, ANOVA, Correlação de Pearson), exigem amostras com um menor número de participantes, comparados a testes não paramétricos (ex: Qui-quadrado, Teste exato de Fisher, Mann-Whitney, Wilcoxon, Correlação de Spearman). Um tutorial<sup>8</sup> sobre a escolha do teste estatístico adequado para o seu estudo pode ser obtido em [www.dentalpress.com.br/bioestatistica](http://www.dentalpress.com.br/bioestatistica).

d) A variabilidade dos dados obtidos. Quanto maior a variabilidade dos dados (desvio-padrão, variância, desvio interquartilico), maior deverá ser o tamanho da amostra. Em geral, amostras pareadas pressupõem um melhor controle da variabilidade dos dados.

Então, uma questão importante surge: será que os pesquisadores da ciência ortodôntica têm, realmente, se preocupado em dimensionar adequadamente as amostras utilizadas em seus estudos? Recentemente, o Instituto de Pesquisa Craniofacial e Odontológica dos Estados Unidos determinou a inclusão de todas essas informações nas pesquisas clínicas candidatas a obter recursos financeiros dessa instituição de fomento<sup>9</sup>.

Embora ausentes na literatura odontológica, alguns estudos têm sido publicados na literatura médica sobre a adequação do tamanho das amostras ou do poder dos testes estatísticos utilizados em pesquisas clínicas. Em geral, a análise de estudos que não conseguem observar diferenças estatisticamente significativas entre um grupo tratado e o controle revela o emprego de amostras com dimensões inadequadas para ler diferenças clinicamente importantes, ou seja, a presença frequente do erro tipo II, falso negativo<sup>2,11</sup>.

Um estudo analisando 2.000 pesquisas clínicas sobre o tratamento da esquizofrenia<sup>10</sup> revelou que apenas 1% das pesquisas publicadas

(n=20) apresentou um estudo do poder estatístico ou cálculo amostral. Os autores observaram, ainda, que apenas 3% dos estudos analisados (n=60) apresentaram um tamanho amostral suficiente para ler uma diferença de 20% entre os resultados das terapias empregadas.

### Erro do método

Outro tipo de erro que pode ocorrer em estudos científicos está relacionado à confiabilidade da leitura dos dados obtidos, ou seja, a capacidade dos mesmos de representar a verdade sobre o fenômeno que ora está sendo examinado e de serem reproduzíveis em um outro momento<sup>4,13</sup>.

Embora comum em estudos com variáveis contínuas (ex: grandezas cefalométricas)<sup>13</sup>, o cálculo de erros de medidas pode e deve, também, ser examinado nos dados não paramétricos (ex: estudos epidemiológicos sobre a prevalência da má oclusão). Entretanto, assim como nem sempre é possível coletar um tamanho apropriada da amostra para um determinado estudo, por questões éticas ou econômicas, em algumas situações é impossível fazer uma análise adequada do erro do método — por exemplo, quando da “perda” inevitável do espécime no momento da leitura dos dados. Frequentemente, essa perda é comum, entre outras circunstâncias, nos ensaios de cisalhamento de braquetes, após o descolamento do acessório, impedindo a repetição do procedimento e obtenção de uma nova mensuração dos níveis de força empregados para a fratura da colagem.

O erro de um determinado método empregado por um pesquisador é conseguido, geralmente, duplicando-se as medidas realizadas em um determinado intervalo de tempo. Dois tipos de erros são geralmente examinados: o erro casual, que mede o grau de precisão das medidas e o erro sistemático, que examina a reprodutibilidade ou replicabilidade das medidas executadas<sup>4</sup>. Uma grande parte da variância de uma variável examinada está relacionada com a imprecisão das medidas obtidas, ou seja, o erro do método, o que

tem, também, implicação no dimensionamento do tamanho da amostra examinada<sup>3</sup>. Quanto menores os erros do método, ou das mensurações, maior a validade ou confiabilidade dos dados obtidos<sup>3</sup>. As medidas devem, portanto, ser “precisas e reproduzíveis”. Define-se como “precisão” a capacidade do método de fazer uma leitura exata do fenômeno examinado; e como reprodutibilidade, a possibilidade dessa medida ser repetida em diferentes momentos.

Apesar de amplamente divulgados e exigidos nos periódicos mais conceituados e de maior fator de impacto, ainda se pode constatar inúmeros artigos científicos sem o adequado exame do erro do método empregado na obtenção dos dados.

### PROPOSIÇÃO

Levando-se em consideração a importância do adequado dimensionamento das amostras analisadas em um estudo científico, assim como a confiabilidade e reprodutibilidade dos dados obtidos a partir dessas amostras, esse estudo tem o objetivo de analisar a frequência com que o emprego do cálculo amostral e do erro do método é relatado na literatura ortodôntica nacional e internacional.

### MATERIAL E MÉTODOS

Dois periódicos científicos relacionados à ciência ortodôntica foram selecionados para a obtenção da amostra do presente estudo. Um publicado em língua portuguesa e editado no Brasil, a *Revista Dental Press de Ortodontia e Ortopedia Facial* (Dental Press Editora, Maringá/PR, Brasil) e outro publicado em língua inglesa e editado nos Estados Unidos, o *American Journal of Orthodontics and Dentofacial Orthopedics*, ou *AJO-DO* (Elsevier, Sant Louis, EUA). Esses dois periódicos foram selecionados, pois representam, segundo a Capes (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), as revistas científicas dessa especialidade com o maior impacto científico no Brasil e no mundo, respectivamente.

A análise dos artigos publicados foi feita, inicialmente, pelo resumo do artigo, sendo observado se o mesmo descrevia alguma análise estatística dos resultados. Em caso positivo, iniciava-se a leitura das seções *Material e Métodos* e *Resultados*, por um único examinador. Uma leitura minuciosa dessas seções foi realizada para se constatar se o artigo publicado apresentava ou não uma análise do cálculo amostral, ou do “poder” do teste estatístico utilizado, e a descrição do erro do método (quando possível de ser aplicado naquele determinado estudo). Não houve qualquer julgamento sobre o grau de adequação dos métodos estatísticos empregados e/ou da confiabilidade dos seus resultados.

Na *Revista Dental Press* foram examinados apenas artigos indexados na seção “artigos inéditos”, enquanto no *AJO-DO* foram analisados somente artigos publicados na seção “artigos originais” (*original articles*). Nessas seções, foram analisados apenas artigos nos quais foram empregados métodos estatísticos descritivos e/ou analíticos. Casos clínicos e revisões de literatura não foram, por razões óbvias, examinados.

### Cálculo amostral

Inicialmente, um estudo piloto foi feito a partir da leitura dos 20 últimos artigos originais publicados no periódico nacional (*Revista Dental Press*) no ano de 2007 (v. 12, n. 4, 5 e 6), com o intuito de se obter a proporção de artigos publicados nessa revista que apresentavam a descrição do cálculo amostral. Nessa busca inicial, apenas 1 artigo, entre os 20 examinados, descrevia, de alguma forma, o cálculo amostral empregado no estudo, o que corresponde a uma proporção de 5%. Em 16 estudos (80%), dentre os 20 analisados, a replicabilidade dos dados foi considerada possível e, portanto, o estudo do erro do método. A análise desses 16 artigos evidenciou o uso do erro do método em 8 trabalhos, ou seja, uma proporção de 50%.

Estabeleceu-se o cálculo amostral através do *software* Bioestat 5.0 (disponível em: [www.mamiraua.org.br](http://www.mamiraua.org.br)), considerando-se o emprego de um

teste estatístico binomial bicaudal, com poder de 80% (0,8) e nível alfa de 5% (0,05) capaz de ler uma diferença de 20% entre os artigos publicados entre as duas revistas. Assim sendo, o tamanho da amostra foi definido como 49 artigos para cada grupo para a análise do cálculo amostral e 93 artigos por grupo para a leitura do emprego do erro do método.

Foram selecionados para leitura os volumes mais recentes das duas revistas, visto representarem a tendência mais atual do formato dos artigos publicados em ambos os periódicos investigados. Com o intuito de se obter o tamanho da amostra predeterminado para a aplicabilidade do cálculo amostral, foi necessária, inicialmente, a leitura de todos os artigos publicados na *Revista Dental Press de Ortodontia e Ortopedia Facial* publicados no ano de 2007 (v. 12), nos números 1 a 6, e em uma edição publicada no ano de 2006 (v. 11, n. 6), totalizando 51 artigos publicados em 7 números da revista. No periódico *AJO-DO* foi necessária, inicialmente, a leitura de 5 números da revista — publicados nos meses de setembro (v. 132, n. 3), outubro (v. 132, n. 4), novembro (v. 132, n. 5) e dezembro de 2007 (v. 132, n. 6) e um número do volume 133, publicado em janeiro de 2008 (n. 1), totalizando o exame de 57 artigos.

Dentre os 51 artigos lidos na *Revista Dental Press* e 57 artigos lidos no *AJO-DO*, foi considerado possível executar o erro do método em 43 dos estudos publicados na *Revista Dental Press* e 41 publicados no *AJO-DO*. Assim, para se alcançar o número determinado no cálculo amostral para se identificar uma diferença na proporção de 20% ( $\alpha = 5\%$ ,  $\beta = 20\%$ ) foi necessária a leitura de mais 6 números da *Revista Dental Press* (v. 11, n. 1 a 5; e v. 10, n. 6,) e 7 números do *AJO-DO* (v. 132, n. 1 e 2; e v. 131, n. 2 a 6), totalizando mais 49 artigos na *Revista Dental Press* (total, n= 92) e 53 no *AJO-DO* (total, n= 94).

### Erro do método

Quinze artigos publicados em cada uma das revistas (n=30) foram aleatoriamente relidos pelo

mesmo examinador, uma semana após a primeira leitura, com o intuito de se avaliar a replicabilidade da avaliação executada. O teste de Kappa, com nível de confiança de 95% ( $\alpha = 0,05$ ), foi utilizado para analisar a replicabilidade do exame. O nível de concordância foi definido tanto para a presença do estudo do cálculo amostral quanto para a presença da análise do erro do método.

### Análise estatística

As frequências observadas na aplicação do cálculo amostral e do erro do método nos dois periódicos examinados foram comparadas através do teste binomial, para duas proporções. O nível de confiança estabelecido foi de 95% ( $\alpha = 0,05$ ).

## RESULTADOS E DISCUSSÃO

### Replicabilidade da leitura dos artigos

A análise do erro do método revelou uma excelente replicabilidade, na releitura dos artigos, para a definição do percentual de emprego do cálculo estatístico, com 100% de concordância (Kappa= 0,9,  $p < 0,001$ , IC95= 1,0–0,53). Quanto à análise da replicabilidade para a descrição do emprego do erro do método nos artigos examinados, foi observada uma concordância de 97% (Kappa= 0,79,  $p < 0,001$ , IC95= 1,0–0,44). Esses resultados preliminares revelam uma excelente reprodutibilidade do método utilizado na leitura dos artigos.

### O emprego do cálculo amostral

O planejamento do tamanho da amostra é frequentemente importante e quase sempre difícil de ser executado, requerendo um cuidado ao eleger os objetivos científicos e ao obter informações apropriadas antes mesmo do início de cada estudo.

Os resultados da presente investigação revelaram que apenas 3,9% dos artigos publicados na *Revista Dental Press*, nos anos de 2005-2007, apresentavam dados sobre o cálculo amostral; esse percentual era significativamente menor ( $p = 0,008$ ) quando comparado ao dos artigos publicados no *AJO-DO*, que apresentavam pouco

mais que 21% dos seus artigos com a descrição dos métodos utilizados para o cálculo amostral (Tab. 1). Esses dados refletem uma desatenção preocupante dos pesquisadores e revisores com esse importante fator de introdução de erros na avaliação estatística dos dados obtidos publicados nesses dois importantes periódicos.

### Estudo do erro do método

No que concerne à preocupação quanto ao emprego do erro do método, os resultados da presente investigação revelaram um quadro diferente daquele apresentado para o uso do cálculo do tamanho da amostra. Na *Revista Dental Press de Ortodontia e Ortopedia Facial*, em apenas pouco mais de 15% dos artigos publicados não era possível, aos autores, duplicar as medidas realizadas no estudo e, dessa forma, efetuar o estudo do erro do método. Para o periódico *AJO-DO*, o estudo do erro do método foi considerado como impossível de ser executado em 28,1% dos artigos publicados (Tab. 2).

Entre os artigos que possibilitavam o estudo do erro do método, a maioria reportava o emprego de algum modelo de análise do erro do método (Tab. 2). Para a *Revista Dental Press*, 60,9% dos artigos (n=56) onde era possível o estudo do erro do método haviam realizado esse tipo de análise; enquanto no *AJO-DO* 76,6% dos artigos (n=72) apresentavam o estudo do erro do método (Tab. 2).

Comparativamente, nos artigos publicados no *AJO-DO* (76,6%), os autores pareciam um pouco mais preocupados com a questão do estudo do erro do método, em comparação aos publicados na *Revista Dental Press* (60,9%/  $p=0,02$ ). Entretanto, ressalta-se o número expressivo de artigos científicos publicados na *Revista Dental Press* que utilizaram alguma ferramenta de controle de erro do método. Deve-se considerar, ainda, o fato dos artigos analisados no *AJO-DO* terem sido publicados nos últimos 12 meses (fev. 2007 a jan. 2008), enquanto na *Revista Dental Press* foi necessária, com o intuito de compor a amostra

TABELA 1 - Frequência absoluta e relativa (%), diferença relativa, valor p e Poder para os artigos com análise estatística (A) e com descrição do cálculo amostral (B) publicados na *Revista Dental Press de Ortodontia e Ortopedia Facial* e no *American Journal of Orthodontics and Dentofacial Orthopedics*.

Descrição	Dental Press 2005-2007	AJO-DO 2007-2008	Diferença (%)	Valor p	Poder
Artigos originais com análise estatística (A)	51 (100%)	57 (100%)	----		
Artigos com descrição do cálculo amostral (B)	2 (3,9%)	12 (21,1%)	-17,2	0,008**	0,78

\*\*  $p<0,01$ .

TABELA 2 - Frequência absoluta e relativa (%), diferença relativa, valor p e Poder para os artigos nos quais seria possível o emprego do erro do método (C) e com a descrição da análise do erro do método (D), publicados na *Revista Dental Press de Ortodontia e Ortopedia Facial* e no *American Journal of Orthodontics and Dentofacial Orthopedics*.

Descrição	Dental Press 2005-2007	AJO-DO 2007-2008	Diferença (%)	Valor p	Poder
Artigos originais com análise estatística (C)	92 (84,3%)	94 (71,9%)	12,4		
Artigos com descrição do cálculo amostral (D)	56 (60,9%)	72 (76,6%)	-14,6	0,02*	0,64

\*  $p<0,05$ .

com um tamanho adequado, a inclusão de artigos publicados nos anos de 2005 e 2006.

Os dados obtidos revelam que os artigos publicados nas duas revistas apresentam um nível de preocupação relativamente pequeno quanto à aplicação de métodos de dimensionamento do tamanho das amostras estudadas. O resultado é preocupante para ambas as revistas, mas é mais crítico para a revista publicada no Brasil: entre 51 artigos analisados, publicados entre o final de 2005 e o final de 2007, apenas 2 (3,9%) faziam referência ao dimensionamento da amostra.

Por outro lado, os dados são mais animadores para o emprego do estudo do erro do método. A maioria dos estudos publicados em ambas as revistas, nos quais essa análise se fazia possível, reportava o emprego do erro do método, embora esse número seja um pouco maior para a revista americana (AJO-DO).

## CONCLUSÃO

A maioria das pesquisas publicadas na *Revista Dental Press de Ortodontia e Ortopedia Facial* e no *American Journal of Orthodontics and Dentofacial Orthopedics*, quando aplicável, emprega alguma forma de análise do erro do método. Porém, apenas em um número muito pequeno desses artigos os autores executam o cálculo do tamanho adequado para as amostras analisadas. Essa proporção, já pequena (21,1%) para a revista editada nos EUA (AJO-DO), é significativamente menor ( $P=0,008$ ) para a revista de Ortodontia editada no Brasil (3,9%). Pesquisadores e o corpo editorial, de ambas as revistas, deveriam direcionar uma maior preocupação para o exame dos erros inerentes à ausência de tais análises na pesquisa científica, em especial os erros inerentes a um dimensionamento inadequado das amostras.

---

## The use of sample size calculation and error of method in researches published in Brazilian and international orthodontic journals

### Abstract

**Introduction:** Reliable sample size and an appropriate analysis of error are important steps to validate the data obtained in a scientific study, in addition to the ethical and economic issues. **Objectives:** To evaluate, quantitatively, how often the researchers of orthodontic science have used the calculation of sample size and evaluated the error of method in studies published in Brazil and in the United States of America. **Methods:** Two major journals, according to Capes (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), were analyzed through a manual search: *Revista Dental Press de Ortodontia e Ortopedia Facial* and the *American Journal of Orthodontics and Dentofacial Orthopedics* (AJO-DO). Only papers published between 2005 and 2008 were examined. **Results:** The results showed that most of the surveys published in both journals used some kind of analysis of the error of method, when this methodology could be applied. On the other hand, only a very small number of articles published in these journals have any description of how sample size was calculated. This proportion was 21.1% for journals published in the United States (AJO-DO), and was significantly lower ( $p=0.008$ ) for the Journal of Orthodontics published in Brazil (3.9%). **Conclusion:** Researchers and the editorial board of both journals should give greater attention to the examination of errors inherent to the absence of such analyses in scientific research, particularly the errors related to the use of an inadequate sample size.

**Keywords:** Biostatistics. Sample size. Error of method.

---

## REFERÊNCIAS

1. Browner WS, Newman TB. Sample size and power based on the population attributable fraction. *Am J Public Health*. 1998;79(9):1289-94.
2. Freiman JA, Chalmers TC, Smith H Jr, Kuebler RR. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. Survey of 71 "negative" trials. *N Engl J Med*. 1978;299(13):690-4.
3. Houston WJ. The analysis of errors in orthodontic measurements. *Am J Orthod*. 1983;83(5):382-90.
4. Hulley SB, Cummings SR, Browner WS, Grady D, Hearst N, Newman TB. *Designing clinical research — an epidemiologic approach*. 2<sup>nd</sup> ed. Baltimore: Williams and Wilkins; 2001.
5. Last JM. Making the dictionary of epidemiology. *Int J Epidemiol*. 1996;25(5):1098-101.
6. Lenth RV. Some practical guidelines for effective sample size determination. *Am Statistician*. 2001;55(3):187-93.
7. MacFarlane TV. Sample size determination for research projects. *J Orthod*. 2003;30(2):99-100.
8. Normando D, Tjäderhane L, Quintão CCA. A PowerPoint-based guide to assist in choosing the suitable statistical test. *Dental Press J Orthod*. 2010;15(1):101-6.
9. National Institute of Dental and Craniofacial Research. Policies and Procedures for Investigator Initiated Clinical Trials. [Cited 2008 Feb 10]. Available from: [www.nidcr.nih.gov/ClinicalTrials/ClinicalTrialsProgram/DataCoordinator.htm](http://www.nidcr.nih.gov/ClinicalTrials/ClinicalTrialsProgram/DataCoordinator.htm).
10. Phillips C. Sample size and power. What is enough? *Semin Orthod*. 2002;8:67-76.
11. Thornley B, Adams C. Content and quality of 2000 controlled trials in schizophrenia over 50 years. *BMJ*. 1998;31(317):1181-4.
12. Torgerson DJ, Miles JN. Simple sample size calculation. *J Eval Clin Pract*. 2007;13(6):952-3.
13. Valladares JVN, Domingues MHMS, Capelozza Filho L. Pesquisa em Ortodontia: bases para a produção e a análise crítica. *Rev Dental Press Ortod Ortop Facial*. 2000;5(4):89-105.

Enviado em: 7 de agosto de 2008  
Revisado e aceito: 31 de agosto de 2008

**Endereço para correspondência**  
David Normando  
Rua Boaventura da Silva, 567-1201  
CEP: 66.055-090 - Belém / PA  
E-mail: davidnor@amazon.com.br