

## A importância do treinamento, reprodutibilidade e calibragem para a qualidade dos estudos.

### *The importance of training, reproducibility and calibration to the quality of the studies.*

Cristiano Susin \*  
Cassiano Kuchenbecker Rösing \*\*

#### RESUMO

A odontologia está passando por uma profunda modificação na sua forma de produzir, utilizar e interpretar o conhecimento científico, sendo que a necessidade de utilizar a melhor evidência possível tanto para o entendimento dos processos físicos e biológicos, quanto para a tomada de decisão clínica, tem despertado interesse no estudo de áreas até então relegadas a segundo plano. Nesse panorama, a capacidade de avaliar a qualidade dos diferentes estudos e de produzi-los com padrão internacional tem ensejado um aumento na procura de conhecimentos sobre metodologia de pesquisa. A forma pela qual os dados são obtidos numa pesquisa e os procedimentos são realizados podem ter influência definitiva na sua capacidade de gerar evidência. Assim, treinamento, reprodutibilidade e calibragem são palavras que devem fazer parte do dia-a-dia de todos envolvidos no processo de criação e utilização do conhecimento.

Palavras-chave: desenho experimental, treinamento, reprodutibilidade, calibragem.

#### UNTERMOS

Odontologia baseada em evidências; reprodutividade, calibragem

#### INTRODUÇÃO

A odontologia, à semelhança das demais áreas biomédicas, está passando por uma profunda modificação na sua forma de produzir, utilizar e interpretar o conhecimento gerado nas diversas pesquisas. A necessidade de utilizar a melhor evidência possível para o entendimento dos processos físicos e biológicos, bem como para a tomada de decisão clínica, tem ensejado uma crescente necessidade de novos conhecimentos por parte de alunos, professores e clínicos. Isso porque, sabidamente, a única forma de manter-se atualizado em relação a novos conhecimentos é através da literatura disponível. Modernamente, a literatura atualizada não se encontra unicamente nos livros-texto, mas, preferentemente, nos periódicos. Entretanto, mudanças de conduta e embasamento da prática clínica devem estar alicerçados em evidência científica de qualidade. Essa nova perspectiva tem atendido pelo nome de Odontologia Baseada em Evidências e tem chamado para si o compromisso de possibilitar o reconhecimento do melhor conhecimento disponível para cada objetivo específico. (5, 10, 14, 16, 18)

Nesse sentido, todos os aspectos metodológicos que possam ser utilizados para melhorar a qualidade dos estudos devem ser observados pelos pesquisadores quando da execução do experimento e, principalmente, pelos leitores quando feita a apreciação dos achados. Entre os princípios gerais de metodologia que podem e devem estar presentes em todo tipo de

estudo estão aqueles que dizem respeito à consistência ou concordância com que os diferentes tratamentos em estudo são executados e os dados são obtidos. Assim, todos os procedimentos executados durante um estudo devem, dentro do possível, ser treinados e sua reprodutibilidade, avaliada.

Os objetivos desses cuidados são os seguintes:

- a) diminuir a variabilidade dos procedimentos e dos dados obtidos;
- b) aumentar a comparabilidade entre tratamentos, exames de diagnóstico, ou outros objetos em estudo;
- c) aumentar a confiabilidade das informações obtidas nos diversos tipos de estudos, aumentando a probabilidade de serem encontradas diferenças e relações, caso elas realmente existam.

A observância adequada desses procedimentos permite, inclusive, que se diminua o tamanho da amostra necessária para demonstrar determinado efeito. (18) Isso é importante para o pesquisador, uma vez que facilita o trabalho e, igualmente, para o leitor, que tem mais capacidade de entender os resultados obtidos bem como de extrapolá-los para a prática clínica.

#### **TREINAMENTO E CONTROLE DA REPRODUTIBILIDADE NA EXECUÇÃO DE PROCEDIMENTOS**

Embora maior enfoque seja dado à mensuração de eventos, os princípios de treina-

mento e controle da reprodutibilidade são aplicáveis à execução de diferentes procedimentos nos estudos. Assim, desde a confecção de corpos de prova até a realização de um procedimento clínico, podem e devem passar por essas fases. Num estudo de materiais dentários, por exemplo, no qual a forma como os corpos de prova são confeccionados pode influenciar os resultados do estudo, um protocolo deve ser escrito e observado durante a obtenção dos mesmos (amostra). A falta de treinamento pode causar o efeito indesejado de que variações inerentes à manipulação dos materiais contribuam de forma decisiva para os resultados do estudo. (7, 18)

Essas considerações assumem importância maior ainda quando, ao invés de um operador, vários estão envolvidos. Imagine-se um estudo no qual o procedimento clínico é realizado por vários indivíduos: essas pessoas podem apresentar padrões absolutamente diferenciados de confecção de uma restauração, de raspagem de um dente, da realização de um procedimento cirúrgico. Essas diferenças podem ser responsáveis por influenciar definitivamente os resultados do estudo.

Assim, quando a execução de procedimentos puder estar sob a influência de características subjetivas dos indivíduos, medidas de precaução devem ser tomadas. A adoção de um protocolo de treinamento, no qual exista a clara definição dos procedimentos a serem realizados e das técnicas a serem utilizadas, permite

\* Cirurgião-dentista, Mestre em Periodontia - ULBRA/Canoas, Professor de Periodontia - ULBRA/Canoas e ULBRA/Cachoeira do Sul.  
\*\* Cirurgião-dentista, Mestre e Doutor em Periodontia - Unesp/Araraquara, Professor Adjunto de Periodontia - UFRGS e ULBRA/Canoas.

que se alcance alto grau de qualidade numa pesquisa. Juntamente com isso, alguma forma de controle de qualidade deve ser implementada, objetivando que os procedimentos mantenham-se comparáveis ao longo do estudo.<sup>(7, 12, 13)</sup>

Uma solução possível para essa questão seria a instituição de um observador externo, que teria, assim, a responsabilidade de zelar pela qualidade dos procedimentos realizados no estudo.<sup>(7, 13)</sup>

Por muito tempo, a odontologia, com vistas a dirimir esse tipo de problema, valeu-se de operadores ou examinadores ditos "experientes". Os novos paradigmas da pesquisa odontológica não aceitam mais unicamente esse tipo de solução. Isso porque a definição de o quanto experiente um indivíduo é passa por critérios extremamente subjetivos e, muitas vezes, pouco científicos. Muitas vezes, alguém muito experiente em determinado procedimento pode não ser suficientemente sistemático no que realiza, o que, com vistas à pesquisa, é problemático. Outras vezes, extrema experiência pode causar descuidos, que são sempre indesejáveis. Portanto, atualmente, medidas de controle do treinamento e da reprodutibilidade de procedimentos são mandatórios para a obtenção de pesquisa de melhor qualidade.

#### TREINAMENTO E CONTROLE DA REPRODUTIBILIDADE NA AFERIÇÃO DE EVENTOS

Um bom estudo começa pela clara definição do evento a ser observado e pela correta escolha do melhor método para aferi-lo. A falha nessas etapas freqüentemente acarreta inconsistências nos resultados. Uma parte do treinamento deve ser feita através da revisão de conceitos ou parâmetros importantes para o estudo.<sup>(11, 12, 13)</sup> A figura 1 ilustra um fluxograma sugestivo de passos que devem ser enfatizados quando da fase de treinamento para a realização de um trabalho de pesquisa, principalmente concernente a medidas de aferição.

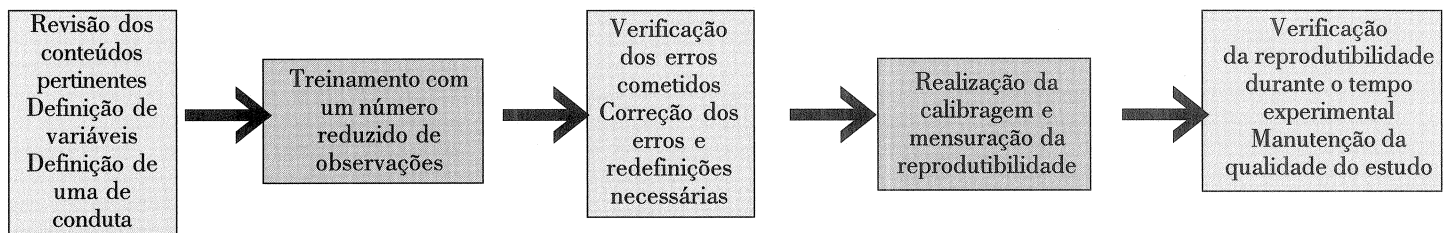


Figura 1 - Fluxograma da fase de treinamento

A criação de um padrão de conduta durante a aferição é conveniente, uma vez que a falta deste pode influenciar negativamente o examinador.<sup>(12, 18)</sup> Por exemplo, quando inexistente um padrão de conduta, pode acontecer que o examinador tome contato com informações importantes para o estudo, em alguns casos, antes do exame e, em outros, somente após. A reafirmação do protocolo de um estudo deve ser realizada durante o estudo para evitar que, no processo de sua execução, existam mudan-

ças de comportamento, o que poderia, inclusive, invalidar os resultados obtidos.

Um erro muito comum durante a fase pré-estudo é o início da aferição da reprodutibilidade previamente a um período de treinamento. Quando esse fato ocorre, normalmente o grau de reprodutibilidade alcançado é baixo, o que acaba por frustrar o pesquisador. Além disso, aos olhos do leitor, que observa baixos níveis de reprodutibilidade do examinador, a qualidade do trabalho decai consideravelmente. Assim, um período de treinamento e de definição de procedimentos é essencial para a obtenção de um grau de reprodutibilidade aceitável no menor espaço de tempo possível. Nesse período, deve-se utilizar um número reduzido de mensurações, sendo que o examinador deve rever os exames e observar em quais condições estão ocorrendo as discrepâncias e qual a melhor forma para contorná-las. Nos estudos que possuem mais de um examinador, o diálogo entre os mesmos é essencial para que se utilizem critérios mais homogêneos de observação.<sup>(11, 12, 13, 18)</sup>

Os procedimentos de tentativa de uniformização de critérios intra e inter-examinadores denominam-se calibragem. A calibragem deve ser realizada em condições o mais próximas possíveis das reais condições do estudo. Esse fato deve ser observado principalmente nos estudos nos quais o que se está mensurando apresenta diferentes graus de severidade ou ocorrência, sendo mais difícil observar diferenças intermediárias do que extremas.<sup>(12, 18)</sup> Não adianta treinar, por exemplo, em uma população sem cáries, a execução do índice CPO-D, se a população-alvo do estudo tem alta prevalência da doença. Provavelmente, se isso for realizado, a reprodutibilidade será boa não devido à consistência do examinador, mas sim devido à inadequação da população em que se realizou a calibragem. Analogias como essa podem ser encontradas em todas as áreas da pesquisa odontológica.

Outro fato muito importante é diminuir o papel da memória do examinador para a segunda aferição. Assim, o número de observações realizadas deve ser grande o suficiente para que o examinador não se lembre do fato nas análises subsequentes. Sabe-se que é mais fácil o examinador lembrar o tipo de condição examinada em um paciente no qual examinou apenas um dente, do que quando a boca toda foi examinada. O desconhecimento do que se está examinando através do cegamento do examinador tam-

bém é uma medida que tende a diminuir a lembrança. Além disso, o tempo entre os exames deve ser o maior possível dentro das possibilidades do estudo. Assim, uma semana entre as mensurações é preferível a um dia, e este, a 15 minutos.

Quando mais de um examinador é utilizado, deve-se eleger um para que sirva de referência. Assim, os outros examinadores devem se adequar às características desse examinador, sendo ele o responsável por intermediar as possíveis diferenças entre os demais componentes.<sup>(11, 13)</sup>

A calibragem deve ser observada como princípio básico de pesquisa em dois níveis: intra-examinador, quando o estudo somente envolve um profissional e intra- e inter-examinador, quando mais de um está envolvido no processo.<sup>(2)</sup>

É importante ressaltar que, embora seja um princípio importante, alguns parâmetros clínicos sofrem de limitações em termos de calibragem, ou seja, não permitem a sua realização<sup>(19)</sup>. Por exemplo, os índices de placa que utilizam a sonda exploradora como forma de verificação da sua presença podem ser chamados de irreversíveis, no sentido de que uma mensuração pode remover a placa e esta pode não ser identificada num segundo momento. Igualmente, índices que revelam a situação no momento do exame não permitem intervalos de tempo muito grandes entre as mensurações, pois elas, certamente, são demonstrativas de condições distintas.

#### REPRODUTIBILIDADE AO LONGO DO TEMPO EXPERIMENTAL

Nos estudos longitudinais, o constante treinamento e a observação da reprodutibilidade ao longo do estudo constituem-se fatores primordiais para que a fase inicial do estudo seja comparável à final. Esse mesmo princípio vale para os estudos de fase experimental longa, sejam eles

laboratoriais, experimentais ou observacionais.<sup>(4, 6, 11, 15)</sup>

Algumas pessoas entendem equivocadamente que, em se tratando de pesquisas sem acompanhamento longitudinal, não existe uma necessidade de controle da reprodutibilidade ao longo do estudo, por essa ter sido realizada previamente. Isso não é verdadeiro nem para os estudos de curta nem para os de longa duração. No primeiro caso, é importante salientar que a atitude do examinador, mesmo que inconscien-

temente, pode mudar da fase de treinamento e calibragem para o decorrer do estudo, ficando este menos atento ou criterioso. Na segunda situação, se o estudo demora um tempo considerável para ser realizado, pode acontecer de o examinador, ao final, ter mudado seu padrão inicial pela aquisição, por exemplo, de maior experiência, pelo desgaste pessoal que esses estudos impingem aos seus realizadores ou, até mesmo, por uma desatenção para os critérios utilizados. Em ambos os casos, o controle da reprodutibilidade é de suma importância para a qualidade do trabalho realizado. Quanto mais longo for um estudo, mais difícil a reprodutibilidade se torna. Entretanto, deve-se, da maneira melhor possível, não perder os referenciais iniciais, permitindo, assim, que os resultados sejam comparáveis. (6, 11, 15)

Nos estudos que visam a comparar dois métodos de diagnóstico, por exemplo, é importante enfatizar que um método que não apresenta boa reprodutibilidade, quando comparado a outro, apresentará pouca concordância, independentemente da qualidade do novo método. Isso é um problema inerente à baixa reprodutibilidade do método e não necessariamente à sua incapacidade diagnóstica. Assim, a mensuração da qualidade da obtenção das observações é imprescindível para uma correta avaliação de um exame diagnóstico. (8, 12)

Sempre que possível, deve-se lançar mão de métodos que possam ser objetivamente reavaliados. Por exemplo, a manutenção de arquivos radiográficos, ou mesmo de modelos de estudo, pode servir de método de conferência ou validação da reprodutibilidade de um exame.

#### **TREINAMENTO E REPRODUTIBILIDADE NOS ESTUDOS MULTICÊNTRICOS**

Os estudos que congregam vários grupos de pesquisa estão se tornando cada vez mais prevalentes. Inclusive, esse tipo de estudo tem recebido maior atenção em termos de capacidade de gerar evidência, justamente pela independência de diferentes centros com relação a determinado evento. Nesses estudos, o controle de reprodutibilidade torna-se mais crítico ainda, pois existem diferenças importantes entre diferentes escolas. A introdução e acompanhamento de todos os princípios discutidos anteriormente deve ser realizada com maior cuidado e preocupação. A elaboração conjunta de um protocolo, com a definição de todas as variáveis que ensejam discussões, e de um código de conduta, são cruciais para a validade do estudo. (4, 6, 7, 13)

Tem sido proposto, para os estudos que são realizados em mais de um centro de pesquisa, que exista uma forma de controle externo às pessoas que estão participando. Essa medida proporcionaria um controle mais efetivo da qualidade e diminuiria possíveis conflitos de interesse ou de posicionamento entre os participantes.

Preferencialmente, essa pessoa não deveria caracterizar vínculos com nenhuma das instituições participantes. (6, 7, 13)

A utilização de equipamentos eletrônicos tem sido uma das formas de permitir melhor reprodutibilidade em todos os tipos de estudo, em especial nos multicêntricos. Entretanto, o simples fato de, em dois centros diferentes, serem utilizados, por exemplo, um eletrodo de flúor, um quantificador de determinados gases, uma sonda periodontal computadorizada e máquinas de ensaio de procedência semelhante, não é indicativo de que exista calibragem interequipamentos. É importante, também, que se calibrem esses instrumentos, pois podem vir a ser fonte de discrepância de resultados.

#### **MENSURAÇÃO DA REPRODUTIBILIDADE ENTRE VARIÁVEIS QUANTITATIVAS**

A reprodutibilidade das mensurações de variáveis quantitativas tem suscitado grande discussão e parece não existir um consenso entre os pesquisadores. Algumas alternativas para a demonstração da concordância podem ser obtidas através do cálculo da média e do desvio-padrão. Assim, se duas mensurações apresentam médias e desvios-padrão compatíveis, seria aceitável supor que são razoavelmente parecidos. Entretanto, essa forma de abordar o problema apresenta limitações importantes, pois duas medidas podem apresentar essas médias muito próximas e serem absolutamente diferentes todo o tempo. A tabela 1 exemplifica esse fato. Essa afirmação é especialmente verdadeira nos casos nos quais a variabilidade dos dados é grande, o que pode, assim, ocultar diferenças. Uma outra forma de se mensurar a reprodutibilidade é através do coeficiente de variabilidade. Entretanto, essa alternativa sofre das mesmas limitações do desvio-padrão. (1, 8, 12)

**TABELA 1 - Demonstrativo das aferições de dois examinadores**

Observações	Observador 1	Observador 2
1	15	13
2	12	15
3	9	10
4	24	27
5	20	16
6	20	15
média	16,66	16,0
± desvio padrão	± 5,15	± 5,29

Fazendo-se uso, ainda, das medidas de variabilidade dos dados, outra possibilidade é o cálculo da média e do desvio-padrão da diferença entre cada um dos pares de observações.

Essas tentativas de se verificar a reprodutibilidade também podem ser realizadas através da confecção de gráficos. Utilizando-se do mesmo princípio de demonstrar as diferenças entre as mensurações, um gráfico poderia ser construído para demonstrar a distância entre cada par de observações. Os gráficos 1 e 2 foram obtidos a partir dos dados encontrados na Tabela 2, como forma de ilustração. No gráfico 1, pode-se observar a diferença entre as observações e comparar sua distância do zero, o que demonstra o grau de concordância entre as observações. Entretanto, no gráfico 2, no qual a diferença entre as observações é confrontada com a média de cada uma das observações, pode-se observar uma informação que o gráfico 1 não apresenta. Com aumento dos valores observados, existe um aumento da diferença entre as mensurações.

**TABELA 2 - Demonstrativo das aferições de dois examinadores**

Observações	Observador 1	Observador 2	diferença entre as observações
1	10	12	-2
2	15	13	2
3	8	9	-1
4	7	7	0
5	6	5	1
6	18	14	4
7	5	4	1
8	13	13	0
9	15	18	-3
10	9	8	1
11	10	12	-2
12	14	18	-4
13	5	7	-2
14	8	6	2
15	20	16	4

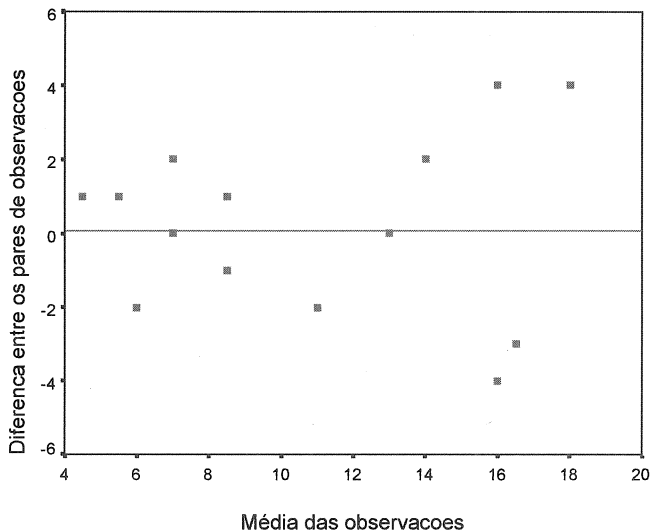


Gráfico 1. Diferença entre as observações e sua distância em relação ao zero

Outras formas têm sido propostas para a mensuração da reprodutibilidade, mas sua discussão foge ao escopo do presente trabalho.

#### MENSURAÇÃO DA REPRODUTIBILIDADE ENTRE VARIÁVEIS CATEGÓRICAS

Muitas das variáveis utilizadas nas pesquisas e mesmo na clínica odontológica são categóricas. Exemplos clássicos são os Índices de Placa e Gingival desenvolvidos nos anos 60 por SILNESS e LÖE. Entretanto, várias formas de classificação de eventos têm sido propostas para qualidade de restaurações, obturações de canal, tipos de oclusão e de perfil, entre outros.

Quadro 1. Classificação de restaurações segundo sua forma anatômica.

Categoria	Descrição
A	a restauração é uma continuação da forma anatômica existente;
B1	a restauração não é contínua com a forma existente - menos de 50% da margem;
B2	a restauração não é contínua com a forma existente - mais de 50% da margem;
C	há perda suficiente de material restaurador para expor a dentina ou a base.

Fonte: Collins et al., 1998.

Quadro 2. Distribuição das observações segundo os critérios de forma anatômica da restauração.

		observação 1				
observação 2	A	A	B1	B2	C	Total
	B1		21	7	2	
B2		7	15	2	1	25
C			9	15	5	29
Total		28	1	2	8	11
			28	32	14	95

Nota: Cálculo do acerto apenas pelo acaso.

$$[(30 \times 28) / 95] + [(25 \times 32) / 95] + [(29 \times 21) / 95] + [(11 \times 14) / 95] = 8,73 + 8,42 + 6,41 + 1,62 = 25,18$$

$$\text{Total} = 25,18 / 25,18 / 95 = 0,26 \quad 0,62 - 0,26 / 1,00 - 0,26 = 0,36 / 0,74 = 0,48 = \text{Kappa}$$

R. Fac. Odontol., Porto Alegre, v. 41, n.1, p. 3 - 7, jul. 2000

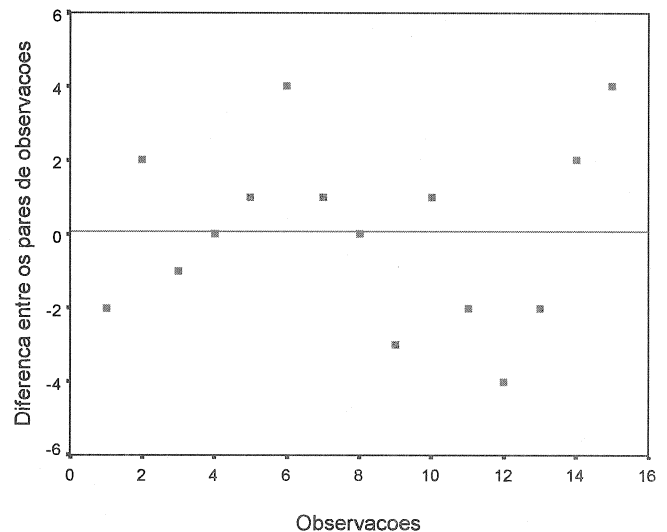


Gráfico 2. Diferença entre as observações confrontada com a média de cada uma das observações.

Quando estudos utilizam esse tipo de variável, a avaliação do grau de concordância também é imprescindível.

O uso de testes como o qui-quadrado e de coeficientes de correlação, bastante freqüente em artigos, não está indicado nesses casos, visto que o objetivo não é testar hipóteses e determinar correlações, mas sim estimar concordância. Para esse fim, o coeficiente Kappa tem sido o mais utilizado. A forma mais fácil de se mensurar a concordância é através da demonstração em percentual<sup>(1, 17)</sup>. Por exemplo, dois observadores utilizando-se dos critérios clínicos de perda de forma anatômica (descritos no Quadro 1) examinaram 95 restaurações. Dentre

essas, 59 observações (62%) apresentaram concordância de critérios, enquanto que, em 32 (33%) das observações, o erro foi de um critério acima ou abaixo. Essa forma de expressar a concordância é útil e de fácil entendimento, entretanto não leva em conta a possibilidade das observações terem sido congruentes apenas pelo acaso (chance). A quantidade de acerto devido ao acaso nessa situação é calculada em 0,26 (cálculo demonstrado junto ao Quadro 2). Dessa forma, a quantidade de acerto devido apenas à concordância entre as observações é dada pelo coeficiente Kappa de 0,48 que, segundo uma tabela de interpretação, é regular (Quadro 3).

Quadro 3. Interpretação do kappa

Valor do kappa	Significado
1	perfeita
0,81 - 0,99	excelente
0,61 - 0,80	bom
0,41 - 0,60	regular
0,21 - 0,40	fraca
0 - 0,20	muito fraca

Fonte: Landis; Koch, (1977).

O coeficiente calculado anteriormente não leva em consideração o grau de diferença entre as observações, sendo chamado de *kappa* não ponderado (*unweighted*). Entretanto, quando existe, conceitualmente, a possibilidade de se admitir que o grau de discordância não é igual entre as categorias, e que errar de B1 para B2 (Quadro 2) é menos importante do que errar de A para C, então deve-se levar em consideração o grau das diferenças. Assim, as diferenças de apenas uma categoria serão consideradas menos sérias do que aquelas de duas ou mais. Esse coeficiente é denominado *kappa* ponderado (*weighted*) e sua aplicação está indicada quando as categorias são ordenadas.

A interpretação do coeficiente *kappa* e dada pelo Quadro 3 e tem aplicabilidade para diferentes parâmetros na área biomédica.

### CONSIDERAÇÕES FINAIS

Tendo por base o exposto no presente artigo, a correta utilização de alguns princípios qualifica um estudo e possibilita uma melhor consistência dos seus achados. Portanto, é de suma importância que os seguintes passos sejam observados quando é feita a preparação, realização e interpretação de trabalhos de pesquisa:

1) elaboração de um protocolo de pesquisa no qual claras definições dos objetos de pesquisa estejam explicitadas;

2) treinamento e determinação da reprodutibilidade previamente ao início do estudo;

3) controle da reprodutibilidade durante o período experimental;

4) observação dos demais princípios básicos da pesquisa, o que torna a evidência claramente delineada sob os paradigmas da ciência vigente.

### SUMMARY

Dentistry is undergoing a deep transformation in its way of producing, using and interpreting the scientific knowledge. The need for utilizing the best possible evidence for the understanding of the physical and biological processes, as well as to clinical decision making, has risen the interest in the study of subjects that were not regarded as important before. In this scene, the capacity of evaluating the quality of different studies and of producing them under international standards has enhanced the seeking for knowledge in scientific methodology. The way by which the data are obtained and the procedures are performed in research might definitively influence its capacity to generate evidence. Thus, training, reproducibility and calibration are principles that have to be part of everything concerning the process of creating and using the knowledge.

### KEYWORDS

Evidence-based dentistry, reproducibility, calibration

### REFERÊNCIAS BIBLIOGRÁFICAS

1. ALTMAN, D.G. **Practical statistics for medical research**. Boca Raton:Chapman & Hall/CRC, 1991. 611p.
2. CHILTON, N.W.; FLEISS, J.L. Design and analysis of plaque and gingivitis clinical trials. *J. Clin. Periodontol.*, Copenhagen, v.13, n.5, p.400-406, May.1986.
3. COLLINS, C.J.; BRYANT, R.W.; HODGE, K.L.V. A clinical evaluation of posterior composite resins restorations: 8-year findings. *J. Dent.*, Guildford, v.26, n.4, p.311-317, May 1998.
4. EATON, K.A.; RIMINI, F.M.; ZAK, E. *et al.* The achievement and maintenance of inter-examiner consistency in the assessment of plaque and gingivitis during a multicentre study based in general dental practices. *J. Clin. Periodontol.*, Copenhagen, v.24, n.3, p.183-188, Mar. 1997.
5. FLETCHER, R.H.; FLETCHER, S.W.; WAGNER, E.H. **Epidemiologia Clínica: elementos essenciais**. trad. Bruce B. Duncan, Maria Inês Schmidt. 3.ed. Porto Alegre : Artes Médicas, 1996. 281p.
6. IMREY, P.B.; CHILTON, N.W. Design and analytic concepts for periodontal clinical trials. *J. Periodontol.*, Chicago, v.63, suplemento. p.1124-1140, Dec. 1992.
7. GOODSON, J.M. Conduct of multicenter trials to test agents for treatment of periodontitis. *J. Periodontol.*, Chicago, v.63, suplemento, p.1058-1063, Dec. 1992.
8. KNAPP, R.G.; MILLER III, M.C. **Clinical epidemiology and biostatistics**. Baltimore: Willians and Wilkins, 1992. 435p.
9. LANDIS, J.R.; KOCH, G.G. The measurement of observer agreement for categorical data. *Biometrics*, Alexandria, v.33, suplemento, n.1, p.159-174, Nov. 1977.
10. NIEDERMAN, R, BADINOVAC, R. Is it philosophy of care or evidence-based dentistry? *Evidence-based dentistry*, London, v.1, n.1, p.3-4, Nov. 1998.
11. PAGE, R.C.; ARMITAGE, G.C., DEROUEN, T.A. *et al.* **Design and conduct of clinical trials of products designed for the prevention, diagnosis and therapy of periodontitis**. Chicago:American Academy Periodontology, 1995. 54p.
12. PEREIRA, M.G. **Epidemiologia: teoria e prática**. Rio de Janeiro:

- Guanabara Koogan, 1995. 583pp.
13. POLSON, A.M. The research team, calibration and quality assurance in clinical trials in periodontics. *Ann. Periodontol.*, Chicago, v.2, n.1, p.75-82, Mar. 1997.
  14. RICHARDS, D.; LAWRENCE, A. Evidence based dentistry. *Br. Dent. J.*, London, v.179, n.7, p.270-273, Nov. 1995.
  15. PLHILSTROM, B.L. Measurement of attachment level in clinical trials: probing methods. *J. Periodontol.*, Chicago, v.63, suppl. 12, p.1072-1077, Dec. 1992.
  16. SACKETT, D.L.; RICHARDSON, W.S.; ROSENBERG, W. *et al.* **Evidence-based medicine. How to practice and teach EBM**. New York:Churchill-Livingstone. 1997. 250p.
  17. SPOLSKY, V.W.; GORNBEIN, J.A. Comparing measures of reliability for indices of gingivitis and plaque. *J. Periodontol.*, Chicago, v.67, n. 9, p.853-859, Sept. 1996.
  18. SUSIN, C.; RÖSING, C.K. **Praticando odontologia baseada em evidências**. Canoas: Ed. Ulbra, 1999, p.35-44.
  19. VAN DER WEIJDEN, G.A.; TIMMERMAN, M.F.; SAXTON, C.A. *et al.* Intra- inter-examiner reproducibility study og gingival bleeding. *J. Periodontol. Res.*, Copenhagen, v.29, n.4, p.236-241, July. 1994.

**Endereço para Correspondência:**  
E-mail: ckrosing@hotmail.com