

INFLUENCE OF IMAGE FILTERS ON THE REPRODUCIBILITY OF MEASUREMENTS OF ALVEOLAR BONE LOSS

INFLUÊNCIA DOS FILTROS DE IMAGEM SOBRE A REPRODUTIBILIDADE DE MEDIDAS DA PERDA ÓSSEA ALVEOLAR

Wivian Eiras Garcia Wong ALVES¹, Evelise ONO², Jefferson Luis Oshiro TANAKA², Edmundo MEDICI FILHO³, Luiz César de MORAES³, Mari Eli Leonelli de MORAES⁴, Julio Cezar de Melo CASTILHO⁵

1- DDS, Undergraduate student, São Paulo State University.

2- DDS, MSc, PhD graduate student, Department of Biosciences and Oral Diagnosis, Sao Paulo State University.

3- DDS, MSc, PhD, Professor & Chair, Department of Biosciences and Oral Diagnosis, Sao Paulo State University.

4- DDS, MSc, PhD, Assistant Professor, Department of Biosciences and Oral Diagnosis, Sao Paulo State University.

5- DDS, MSc, PhD, Associate Professor, Department of Biosciences and Oral Diagnosis, Sao Paulo State University.

Corresponding address: Prof. Titular Edmundo Medici Filho - Faculdade de Odontologia de São José dos Campos - Departamento de Diagnóstico e Cirurgia - Av. Eng Francisco José Longo, 777 - 12245-000 - São José dos Campos - SP - Brasil - Phone: (12) 3947-9054 - Fax: (12) 3947-9010. e-mail: evelise.ono@uol.com.br - edmundo@fosjc.unesp.br

Received: April 18, 2006 - Modification: July 18, 2006 - Accepted: September 05, 2006

ABSTRACT

The reproducibility of measurements of alveolar bone loss on radiographs may be a problem on epidemiologic studies, as they are based on comparisons of the diagnosis of various examiners. The aim of the present research paper was to assess the inter- and intra-examiner reproducibility of measurements of the interproximal alveolar bone loss on non-manipulated digital radiographs and after the application of image filters. Five Oral Radiologists measured the distance between the cemento-enamel junction (CEJ) to the alveolar crest or to the deepest point of the bony defect on 12 interproximal digital radiographs of molars and bicuspids of a dry human skull. The digital manipulation and the linear measurements were obtained with the Trophy Windows software (Trophy®). For each image, six different versions were created: 1) non-manipulated; 2) bright-contrast adjustment; 3) negative; 4) negative with brightness-contrast adjustment; 5) pseudo-colored; 6) pseudo-colored with brightness-contrast adjustment. In order to prevent interpretation bias because of the repetition of measurements, the examiners measured the radiographs in a random sequence. The two-way ANOVA test at 5% level of significance to compare the means of readings of the same operator with each filter indicated $p < 0.05$ for the majority of operators, while the comparison between the mean values of operators using the same filter indicated $p > 0.05$ for all filters. Based on the results, we concluded that linear measurements of interproximal alveolar bone loss on digital radiographs are highly reproducible among examiners. Nevertheless, the application of image filters significantly influenced the degree of intra-examiner reproducibility. Some filters even reduced the reproducibility of intra-examiner readings.

Uniterms: Digital radiography; Periodontal disease; Linear measurements.

RESUMO

A reprodutibilidade de medidas da perda óssea alveolar em radiografias pode se tornar um problema em levantamentos epidemiológicos, que são baseados em comparações de diagnóstico de vários observadores. O objetivo neste estudo foi avaliar a reprodutibilidade intra e inter-examinador de medidas da perda óssea interproximal, em radiografias digitais não manipuladas e após a aplicação de diferentes filtros de imagem. Cinco radiologistas avaliaram a distância da junção cimento-esmalte até a crista óssea alveolar ou ponto mais apical do defeito ósseo em 12 radiografias digitais interproximais da região de molares e pré-molares de crânio humano macerado. As manipulações digitais e as medidas lineares foram realizadas no programa de computador do sistema radiográfico (Trophy®). Para cada imagem foram criadas seis diferentes versões: 1) não manipulada digitalmente; 2) manipulada por ajuste de brilho e contraste; 3) negativo; 4) negativo com ajuste de brilho e contraste; 5) pseudo-colorida; 6) pseudo-colorida com ajuste de brilho e contraste. Para prevenir tendências da mesma leitura pela avaliação repetida das radiografias, determinou-se uma ordem para a interpretação das imagens. O teste ANOVA fator duplo sem repetição, com nível de significância de 5%, realizado para comparação das médias de um mesmo examinador nas diferentes versões de imagem revelou $p < 0,05$ para a maioria dos examinadores, enquanto para análise entre os examinadores, com cada ferramenta, $p > 0,05$. A análise dos resultados nos permite concluir que medidas lineares da perda óssea alveolar interproximal em radiografias digitais apresentam alto grau de reprodutibilidade inter-examinadores, entretanto a aplicação de filtros de imagem não influenciou significativamente no grau de reprodutibilidade dessas medidas. Ao contrário, certas ferramentas reduziram a reprodutibilidade das medidas intra-examinador.

Unitermos: Radiografia digital; Doença periodontal; Medidas lineares.

INTRODUCTION

Periodontal disease may present a wide variation of its appearance, which brings lack of uniformity on the criteria that define its aspects. This is a problem for epidemiologic researches, since they require reliable information¹⁷.

Conventional dental radiography, widely employed as a tool to determine the situation (presence and extension) of periodontal disease, is considered adequate to surveys of periodontal bony loss, although it may underestimate the actual alveolar bone loss. Despite the fact that progresses on the standardization of radiographs have been achieved, their use to qualify the periodontal disease still depends on subjective interpretation. Therefore, measurements of alveolar bone loss have a low level of reproducibility, mainly because of the difficulty to visualize the cemento-enamel junctions (CEJs) and alveolar crests (AC)^{1,4,6,16,17}.

According to some authors^{7,11,12,19}, digital manipulation of the radiographic images may improve the reproducibility and the diagnostic value of radiographic interpretation, as the structures are better visualized. Once the radiographs are digitalized, filters may be readily applied to enhance visualization of the bone architecture for the measurements of bone loss. Researches on density, contrast and edge enhancement have already proved their usefulness to the diagnosis of caries and alterations of the alveolar bone^{15,18}.

In 1990, Hildebolt, et al.¹⁰ assessed the measurements of alveolar bone loss of human skulls on digitized and manipulated interproximal radiographs. The measurements were done by two examiners with more than 15 years of clinical experience. The author did not observe remarkable differences between the readings of the examiners. He concluded that the digital manipulation technology is highly trustable to establish indices of morbidity of the periodontal disease.

According to Akesson, et al.¹ (1992), the interpretation of radiographs may be influenced by their quality, the method of diagnosis, and by the operator. Therefore, they studied the influence of different observers on the accuracy of measurements of the bone level using three radiographic methods. A systematic and substantial inter-examiner variation was observed, even with experienced examiners who received basic trainings from the same school. This demonstrates the importance of relying the exams and the diagnostic comparisons on the evaluations of various examiners.

In 1999, Eickholz, et al.⁵ assessed the measurements of alveolar bone loss and the depth of the bony defect on non-manipulated and manipulated digitized interproximal radiographs. Seven different types of image filters were used. They concluded that the digital manipulation did not improve the efficacy of the evaluation of bone loss when compared to the non-manipulated images. They concluded that two of the filters reduced the validity of measurements.

After comparing the image quality of six digital radiography systems, Borg, et al.² (2000) observed that manipulation of images by the histogram did not improve their quality.

Wolf, et al.¹⁹ (2001) evaluated the reproducibility of linear measurements of interproximal bone loss on digitized radiographs after application of several methods of magnification and manipulation. They observed that the agreement of identification of the apical aspect of bone defects was lower than the identification of alveolar crests with non-manipulated images. The intra-examiner reproducibility for assessment of the distance between the CEJ and the alveolar crest of examiners was different, being higher for the most experienced one. In general, the intra- and inter-examiner reproducibility of linear measurements on radiographic images was not increased by the manipulation tools of the study (grayscale manipulation and magnification). On the contrary, under certain circumstances, these tools increased the variability of measurements.

As stated by Mol¹⁴ (2004), and oppositely to what manufacturers say, the digital images do not improve the efficacy of diagnosis if compared to conventional intraoral images. According to the author, there is neither improvement nor reduction of the diagnosis efficacy. The film-based images continue to be technically equivalent, if not superior to the digital images.

Based on the previous information, the aim of the present research paper was to assess the inter- and intra-examiner reproducibility of linear measurements of interproximal bone losses on digital images that were not manipulated and images that suffered different types of manipulation (brightness-contrast adjustment, negative and pseudo-color).

MATERIALS AND METHODS

Twelve interproximal radiographs of the molar and bicuspid regions of both sides of a dry human skull were used in this study. Since the image receptor is smaller than a conventional number 2 intraoral film, each tooth of these two groups of teeth and their antagonists were radiographed separately. A film holder (*Rinn Corporation, Elgin, IL, USA*), adapted to the size of the sensor was used to standardize the radiographic images.

An X-ray machine (*Gendex 765DC - Gendex Dental Systems, Milan, Italy*), operating at 65 kVp and 7 mA, and a digital radiography system (*RadioVisioGraphy - Trophy Radiology, Toulouse, France*) were used. The exposure time for all radiographs was 0.02 seconds. This time was determined on a pilot study in which the best exposure time would be the one that provided the best image contrast.

All images were saved on non-compressed TIFF (*Tagged Image File Format*) format and analyzed by 5 examiners, all Oral Radiologists, on a computer with a flat, 17-inch screen (*LG Studioworks 710E - LG Electronics Corp., Taubaté - Brazil*). The training for utilization of image filters was given by the same institution for all examiners. The examiners had to identify the distance from the cemento-enamel junction (CEJ) to the alveolar crest (AC) or to the deepest extension of the bony defect (BD).

The digital manipulations and the linear measurements

from the CEJ to the AC or from the CEJ to the BD were done on the Trophy Windows (RadioVisioGraphy – *Trophy Radiology, Toulouse, France*) software which is part of the digital radiographic system. The magnification of images was the same for all measurements. For each image, six different versions were created: 1) non-manipulated; 2) brightness-contrast adjustment; 3) negative; 4) negative with brightness-contrast adjustment; 5) pseudo-colored; 6) pseudo-colored with brightness-contrast adjustment.

The distances that represent the mesial and distal bone loss of the mandibular first and second bicuspsids, and of the mandibular first and second molars were measured by the five examiners on each version of the twelve images. A total of 48 images were measured. Each measurement was redone after a two-week interval. To prevent bias of the same measurement, the authors established an order for interpretation of images. Each examiner measured the non-manipulated version of one image, then the brightness-contrast version of a second image, then the negative version of a third image, and so on. The images were randomly selected for measurements.

RESULTS

In order to check the error of the method, the measurements from the first and second readings of each examiner were analyzed by linear regression analysis. As long as there were no significant errors in the method ($r^2=0.9$), that means, differences between the first and the second readings were not significant, the mean between the first and the second measurements could be used for statistical analysis.

In order to verify if there were any significant differences among the mean values of the measurements of one same examiner with each filter, data were analyzed using two-way ANOVA test (5% level). Although the mean values of examiners could be equivalent, the measurements of each

region could vary according to each examiner. This could happen because of the higher or lesser difficulty of each observer to measure a certain region with aid of one or another image filter. This was the reason why the authors decided to use the two-way ANOVA. In other words, with this statistical analysis, it would be possible to determine if one specific image filter would facilitate the reading of one specific region. Therefore, it was necessary to determine the influence of each region over the mean values of the measurements of each examiner.

Table 1 displays the mean and standard deviation (SD) of the measurements taken by each examiner applying each image filter. Table 2 presents the *p*-value for the comparisons among filters and the *p*-value for the interaction between the variable “region” and the variable “filter”.

The one-way ANOVA test was used to compare the measurements of examiners using one specific image filter. In this case, it was not necessary to verify the influence of the variable “region” over the variable “examiner”.

Table 3 shows the mean and SD of the measurements of the 5 examiners using the same filter. Table 4 demonstrates the *p*-value of the comparison among means of measurements of each examiner using one specific image filter.

DISCUSSION

The periodontal disease is clinically diagnosed by assessment of the probing depth and clinical attachment level, which ranges between 1 and 3 millimeters. However, radiographs are usually used on the evaluation of disease.

The radiographic examination is a complement to clinical diagnosis, since they provide parameters for an adequate treatment planning for each case and helps in the evaluation of alveolar bone response to treatment. Nevertheless, one of the limitations of radiographs is the fact that they are 2D representations of 3D structures and therefore

TABLE 1- Mean and standard deviation (SD) for the values of each examiner using different image filters

EXAM	IMAGE FILTER					
	A Mean±SD	B Mean±SD	C Mean±SD	D Mean±SD	E Mean±SD	F Mean±SD
1	2.079±1.36	1.975±1.34	2.146±1.37	2.088±1.33	1.971±1.19	2.075±1.34
2	2.250±1.37	2.233±1.41	2.329±1.43	2.321±1.40	2.425±1.42	2.400±1.42
3	2.483±1.38	2.479±1.40	2.496±1.36	2.529±1.37	2.679±1.41	2.713±1.36
4	1.950±1.21	1.867±1.22	1.971±1.11	2.079±1.25	2.013±1.30	2.088±1.15
5	2.483±1.49	2.596±1.51	2.479±1.49	2.496±1.51	2.629±1.42	2.600±1.50

- A – Non-manipulated image
- B – Brightness-contrast adjustment
- C – Negative
- D – Negative with brightness-contrast adjustment
- E – Pseudo-color
- F – Pseudo-color with brightness-contrast adjustment

superimposition of images that may mask the actual condition of the alveolar bone may occur^{8,9,12}.

The alveolar bone loss is the main characteristic of the periodontal disease. The height of the alveolar bone may be determined either surgically or in a non-invasive manner by radiographs. However, there is a tendency of underestimation of bone loss with the non-invasive method^{1,4,6,17}.

Morbidity indices based on radiographs are also limited by the variations of densities of the oral structures, which reduces the accuracy of differentiation of the buccal and the lingual alveolar bone heights, especially when the alveolar crests have lower density or are superimposed by the teeth. In the presence of a non-treated intra-bony defect,

it is more difficult to find its apical portion because the bone structure is diffuse. That may be the explanation for the high prevalence of underestimation or overestimation of bone loss during measurements of distances from the CEJ to the BD¹⁰.

Such obstacles may bring differences on the judgments of the examiner, and are a problem for comparisons among distinct epidemiological surveys¹⁷.

Therefore, for many authors^{3,5,7} the digital radiography has many favorable characteristics for the diagnostic procedure when compared to conventional radiography. One of them is the possibility to make digital adjustments to increase the image quality, which facilitates the detection of

TABLE 2- *p-value* for the comparisons among the values of each image filter of each examiner and the *p-value* for the interaction between the regions and the filters

	EXAMINER				
	1	2	3	4	5
<i>p-value</i> Filters*	0.007	0.017	0.014	0.003	0.155
<i>p-value</i> Interactions**	0.236	0.998	0.786	0.318	0.245

* $p < 0,05$ – Significant intra-examiner differences.

** $p < 0,05$ – Influence of the region on the results of the readings by each filter.

TABLE 3- Mean and SD of the values of the examiners using the same filter

FILTER	EXAMINER				
	1 Mean±SD	2 Mean±SD	3 Mean±SD	4 Mean±SD	5 Mean±SD
A	1.975±1.341	2.233±1.417	2.583±1.510	1.867±1.224	2.479±1.409
B	2.146±1.378	2.329±1.436	2.446±1.498	1.971±1.199	2.496±1.367
C	2.088±1.338	2.321±1.407	2.442±1.530	2.079±1.251	2.529±1.376
D	1.983±1.338	2.425±1.407	2.671±1.530	2.013±1.251	2.679±1.376
E	1.983±1.244	2.425±1.427	2.671±1.426	2.013±1.306	2.679±1.414
F	2.117±1.295	2.400±1.421	2.613±1.508	2.088±1.154	2.713±1.36

A – Non-manipulated image

B – Brightness-contrast adjustment

C – Negative

D – Negative with brightness-contrast adjustment

E – Pseudo-color

F – Pseudo-color with brightness-contrast adjustment

TABLE 4- *p-value* of the comparisons among the values of the examiners using the same filter

	FILTER					
	A	B	C	D	E	F
<i>p-value</i> *	0.843	0.662	0.876	0.901	0.563	0.719

* $p > 0,05$ – No significant inter-examiner differences for each filter.

structures on the image. A number of softwares offer a tool to adjust the gray scale, brightness, contrast, edge sharpening, color modification and grayscale inversion to improve visualization by examiners.

However, some authors like Kullendorff and Nilsson¹³ (1996), Eickholz, et al.⁵ (1999), Borg, et al.² (2000), Wolf, et al.¹⁹ (2001) and Mol¹⁴ (2004) stated that the image filters do not influence the diagnostic process.

Considering the importance of reproducibility of measurements of bone loss on epidemiological surveys of periodontal disease, the aims of the present research were to verify the intra- and inter-examiner differences regarding linear measurements of periodontal disease on digital radiographs and if the application of image filters affects this reproducibility.

The linear regression analysis did not demonstrate significant differences between the first and second readings of examiners. Comparing the six measurements of the same examiner (Table 2), each corresponding to one version of the images, the readings of four out of the five examiners were significantly different ($p < 0,05$). In general, the use of image filters did not increase the reproducibility of measurements of the examiners. This result was also found by Wolf, et al.¹⁹ (2001). Nevertheless, these authors stated that the lack of experience of one of the observers of their study was the cause of the low reproducibility of measurements, since the most experienced observer presented a high reproducibility level. In the present study, the measurements of the most experienced observers (1, 2 e 4) were not reproducible ($p < 0,05$).

The inter-examiner results were compared (Table 4) and the p -values demonstrated high level of reproducibility among them using the same image filter ($p > 0,05$). It is clear that even with the non-manipulated image (Filter A), a high p -value was observed ($p = 0,843$). Nonetheless, the filters B (brightness/contrast) and E (pseudo-color) reduced the reproducibility level of measurements.

The results of statistical analyses indicated that application of certain types of image filters affect the radiographic interpretation. Nevertheless, since it was not the aim of the present paper to analyze the validity of measurements, we cannot state that the filters improved the diagnostic process. In this study, it was observed that the filters do negatively affect the readings of the same operator.

Analyzing the measurements of the five examiners, it was observed that the negative filter and negative with brightness/contrast adjustment provided better results, since the p -values for those comparisons were higher. However, we found $p > 0,05$ for all image filters. Therefore, in general there was no improvement in inter-examiner reproducibility of measurements of alveolar bone loss with the application of image filters. It is possible to state, as did Akesson, et al.¹ (1992), that the radiographic interpretation was influenced not only by the radiographs or the method of diagnosis, but also by the examiner. Reliable comparisons of intra-examiner results can be obtained with an index of well-established criteria. This will also allow reliable comparisons of inter-examiner results¹⁷.

CONCLUSIONS

Based on the results, we concluded that:

- In general, the use of image filters did not improve the reproducibility of intra-examiner results. Four out of five examiners presented statistically significant differences ($p < 0,05$) among their measurements regarding the six image filters;
- There were no significant differences among the mean values of measurements of examiners using the same image filter. Therefore, a high level of inter-examiner reproducibility was found, independently from the image filter;
- The negative filter and the negative with brightness/contrast adjustment filter provided higher p -values to the inter-examiner reproducibility.

REFERENCES

- 1- Akesson L, Hakansson J, Rohlin M. Comparison of panoramic and intraoral radiography and pocket probing for the measurement for the marginal bone level. *J Clin Periodontol.* 1992;19:326-32.
- 2- Borg E, Attaelmanan A, Gröndahl HG. Subjective image quality of solid-state and photostimulable phosphor systems for digital intraoral radiography. *Dentomaxillofac Radiol.* 2000;29:70-5.
- 3- Clasen NF, Aun CE. Radiografia digital direta: análise dos principais sistemas e recursos. *Rev Odontol UNICID.* 1998;10(2):109-20.
- 4- Eickholz P, Kim TS, Benn DK, Staehle HJ. Validity of radiographic measurement of interproximal bone loss. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 1998;85(1):99-106.
- 5- Eickholz P, Riess T, Lenhard M, Hassfeld S. Digital radiography of interproximal bone loss; validity of different filters. *J Clin Periodontol.* 1999;26:294-300.
- 6- Eickholz P, Hausmann E. Accuracy of radiographic assessment of interproximal bone loss in intrabony defects using linear measurements. *Eur J Oral Sci.* 2000;108:70-3.
- 7- Guneri P, Lomçali G, Boyacioglu H, Kendir S. The effects of incremental brightness and contrast adjustments on radiographic data: a quantitative study. *Dentomaxillofac Radiol.* 2005;34:20-7.
- 8- Hausmann E. Radiographic and digital imaging in periodontal practice. *J Periodontol.* 2000;71:497-503.
- 9- Hausmann E, Allen K, Christersson L, Genco RJ. Effect of x-ray beam vertical angulation on radiographic alveolar crest level measurement. *J Periodont Res.* 1989;24:8-19.
- 10- Hildebolt CF, Vannier MW, Shrout MK, Pilgram TK, Province M, Vahey EP, et al. Periodontal disease morbidity quantification. II. Validation of alveolar bone loss measurements and vertical defect diagnosis from digital bite-wing images. *J Periodontol.* 1990;61:623-32.
- 11- Hildebolt CF, Pilgram TK, Yokoyama-Crothers N, Fletcher G, Helbig JL, Barlett TQ, et al. Reliability of linear alveolar bone loss measurements of mandibular posterior teeth from digitized bitewing radiographs. *J Clin Periodontol.* 1998;25:850-6.
- 12- Jeffcoat MK, Wang IC, Reddy MS. Radiographic diagnosis in periodontics. *Periodontol* 2000. 1995;7:54-68.

- 13- Kullendorff B, Nilsson M. Diagnostic accuracy of direct digital dental radiography for the detection of periapical bone lesions. II. Effects on diagnostic accuracy after application of image processing. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 1996;82(5):585-9.
- 14- Mol A. Imaging methods in periodontology. *Periodontol* 2000. 2004;34:34-48.
- 15- Nair MK, Ludlow JB, May KN, Nair UP, Johnson MP, Close JM. Diagnostic accuracy of intraoral film and direct images for detection of simulates recurrent decay. *Oper Dent.* 2001;26(3):223-30.
- 16- Pepelassi EA, Diamanti-Kipioti A. Selection of the most accurate method of conventional radiography for the assessment of periodontal osseous destruction. *J Clin Periodontol.* 1997;24:557-67.
- 17- Theilade J. An evaluation of the reliability of radiographs in the measurements of bone loss in periodontal disease. *J Periodontol.* 1960;31:143-53.
- 18- Wenzel A, Borg E, Hintze H, Grondahl HG. Accuracy of caries diagnosis in digital images from charge-coupled device and storage phosphor systems: an in vitro study. *Dentomaxillofac Radiol.* 1995;24(4):250-4.
- 19- Wolf B, Von Bethlenfalvy E, Hassfeld S, Staehle HJ, Eickholz P. Reliability of assessing interproximal bone loss by digital radiography: intrabony defects. *J Clin Periodontol.* 2001;28:869-78.