ORIGINAL ARTICLE

RN Smith
DL Lath
A Rawlinson
M Karmo
AH Brook

Gingival inflammation assessment by image analysis: measurement and validation

Authors' affiliations:
RN Smith, *M Karmo*, *AH Brook*, School of Dental Sciences, University of Liverpool, Edwards Building, Liverpool, UK
DL Lath, Department of Oral Health and Development, Claremont Crescent, Sheffield, UK
A Rawlinson, Department of Adult Dental Care, University of Sheffield, School of Clinical Dentistry, Claremont Crescent, Sheffield, UK

Correspondence to:
Darren L Lath
Department of Oral Health and Development
School of Clinical Dentistry
University of Sheffield
Claremont Crescent
Sheffield S10 2TA
UK
Tel.: +44 114 2717887
Fax: +44 114 2717843
E-mail: d.lath@sheffield.ac.uk

Dates:
Accepted 1 October 2007

To cite this article:
Int J Dent Hygiene **6**, 2008; 137–142
Smith RN, Lath DL, Rawlinson A, Karmo M, Brook AH. Gingival inflammation assessment by image analysis: measurement and validation.

© 2008 The Authors.
Journal compilation © 2008 Blackwell Munksgaard

Abstract: *Aim:* Gingival inflammation may be caused by injury or plaque-related diseases and reduction in inflammation can be a useful indicator of gingival recovery. There has been little research on development of non-index methods to measure gingival condition. The aims of the study were to investigate the reliability of the measurement of changes in gingival redness and swelling, using image analysis, and to compare this approach with an established method for assessing gingival overgrowth [J Clin Periodontol 28 (2001) 81]. *Method:* Twenty volunteers with gingival inflammation were recruited and digital images were taken. Duplicate measurements were made on the first visit by two examiners. At a subsequent visit following periodontal treatment, second images were taken. Gingival changes were determined by assessing redness and tooth surface area visible between the level of the inter-proximal papillae and the gingival margin. Tooth area measurements were compared with the established gingival overgrowth method. *Results:* The method showed excellent reliability for both intra- and inter-examiner measurements of 0.968–0.998 and 0.769–0.947, respectively, according to the classification by Donner and Eliasziw of the Fleiss coefficient of reliability (repeat measures taken during the patients' first attendance). High correlation was found for gingival encroachment when compared with the established gingival overgrowth method. *Conclusion:* This technique proved a reliable method for investigating changes in gingival redness. High correlation was found for gingival encroachment when compared with an established method.

Key words: gingiva; image analysis; inflammation

Introduction

Until the 1950s, gingival health was only assessed subjectively as good, medium or poor. Many indices have been described since then, such as the P–M–A index (1), which was later modified (2). This method relies on recording the location of inflammation in the papilla (P) or gingival margins (M) and noting areas of attached gingivae affected (A). The WHO promoted the development of the Periodontal Index which was used widely up to the later part of the 20th century (3). This index was the first used in longitudinal studies to show that scores lowered after treatment as the gingival inflammation index scoring was reversible.

Many current methods are derivatives of earlier existing systems such as Harris and Ewart's 0–3 system (4). Scores of 0–3 define increasing degrees of gingival encroachment over the tooth surface and severity. Methods, such as the photographic system produced by Ellis *et al.* (5), used photographs of patient's teeth that were projected onto a large screen. These were assigned a score indicating the encroachment level of the anterior papilla from 0 to 3. This method was applied to the labial surfaces of the anterior teeth. The scoring system was based on that developed by Seymour *et al.* (6) whereby the score 0 indicated no health problems, 1 defined slight granulation of appearance of the papilla, 2 indicated slight gingival encroachment over the tooth surface less than a quarter of the tooth width and 3 described severe encroachment over a quarter of the tooth width. This method used basic image analysis, but relied upon the three-stage scoring applied to each of 10 papillae regions (five upper and five lower). The limitations of most indices are related to the subjectivity of examiner scoring and the level of method standardization achievable.

This study aimed to investigate a method of measuring inflammatory changes in the gingival tissues that may be used to study the effects of treatment on individual patients and treatment outcomes in clinical trials, and to compare this approach with an established method for assessing gingival overgrowth (5).

Method

Patients

Twenty volunteers gave their written, informed consent to take part in this study conforming to the South Sheffield Ethics Committee regulations. The patients were chosen from those attending the Periodontology and Dental Hygiene clinics, Charles Clifford Dental Hospital, Sheffield. The assessors (RNS and DLL) were informed when a suitable patient attended clinic with apparent gingival inflammation. The causal factors were not necessarily known at this stage. Each patient was seen on two occasions (pre- and post-treatments) and digital photographs were taken on each visit. Both assessors took two sets of images each on the first visit for reliability calculations. The second visit was not less than 1 month after the first and during this time non-surgical periodontal treatment was provided by one of the Hygienists (the post-treatment assessment was undertaken by RS alone as repeat measurements were not required). This consisted of advice on plaque control, scaling and root surface debridement.

Imaging system

The acquisition apparatus described by Smith *et al.* (7, 8) includes a frame designed by the first author and constructed within the Department of Oral Health and Development, School of Clinical Dentistry, Sheffield. This frame rotates around a Cephalometric head positioning apparatus. It has a platform mounting for a 32-bit Kodak Nikon DCS410 Digital Camera [CCD Dynamic Random Access Memory (DRAM) imager (Eastman Kodak, Rochester, NY, USA), giving an ISO of 100, providing 1.5 mega pixel resolution in an array of 1012 × 1524 pixels, producing 4.6 MB TIF files], with a 90-mm high-quality Elicar macro lens. The camera position can be adjusted and recorded for height and in forward/backward position to accommodate different facial sizes. The frame also supports two Portaflash 220 slave flashguns (Jessops, Leicester, UK) with white opacity filters to soften illumination. Each light is covered by polarizing film, with the polarizing effect direction set the same for both flashes and at 90° to a circular polarizing filter attached to the camera lens. A flashgun on the camera triggers the slave flashes. The trigger flash is covered with exposed film so that only the infra-red light required to initiate the slave flashes is transmitted, without affecting the controlled light supply. The whole frame can be rotated around the patient's head until correct alignment with the tooth of interest is obtained. Parallax was not a consideration as all repeats were taken under the same conditions.

Image analysis

After image acquisition using Adobe Photoshop (V5.02; Adobe Systems Ltd, San Jose, CA, USA), images were measured as shown in Figs 2 and 3. A thresholding process automatically selected a predetermined range of pixel values of red, from the

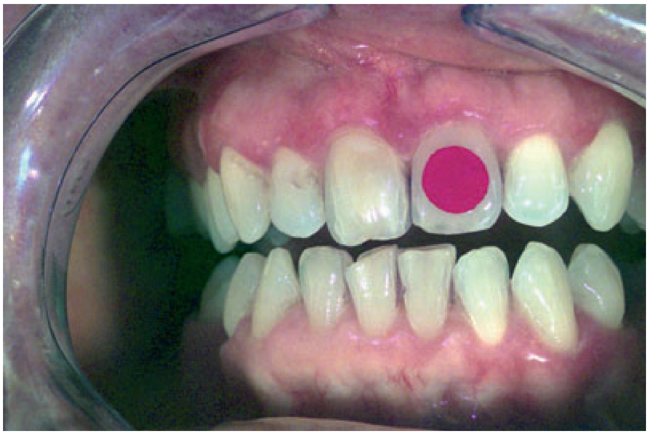

Fig. 1. Image of the anterior teeth of a patient with gingival inflammation showing a red calibration disc attached to the upper left central incisor.

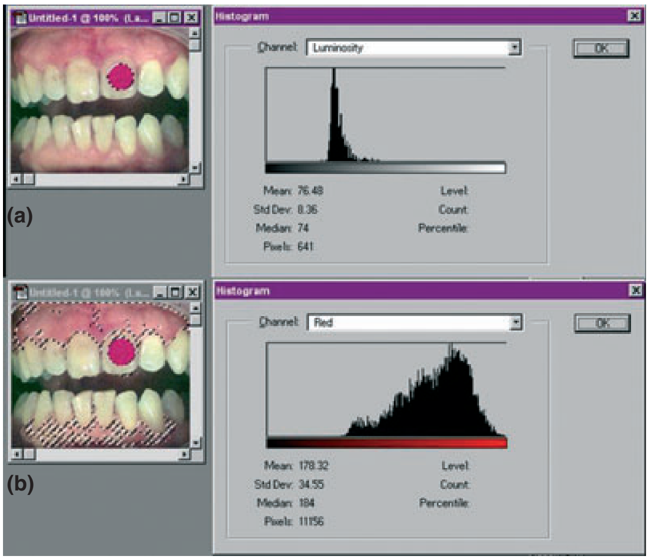

Fig. 2. (a) Image of the isolated red calibration disc and mean red pixel output menu; (b) image of the visible gingival area isolation by thresholding and area measurement is shown in the output menu after calibration.

total 256 available. This process was used to isolate a red disc of known size (Fig. 1, incorporated in the images for colour and size calibration) and then the total gingivae visible on the image: the histogram option was used to give the mean red pixel value of each (Fig. 2) which was then recorded for later subtraction from the red disc value. Calibration and analysis of images was carried out using Image Pro Plus software (V4.0; Media Cybernetics, Silver Spring, MD, USA). This process was repeated on the post-treatment images gained at the patients' subsequent visits. Any slight differences in the image illumination or in patient position would be accounted for by always subtracting the standard red-coloured disc mean red pixel value from that of the gingivae, as the mean value of the red disc should not change in ideal conditions. This method ensures that changes seen are due to gingival changes not random or systematic errors.

All the red discs were cut from the same sheet of red articulating paper, using a sterilized hole punch and were 6 mm in diameter. The discs gave almost identical red pixel values when checked before use. One was attached using the patient's saliva, to the upper central left incisor to facilitate calibration in mm^2 of the labial surface of the upper central incisors that is bounded by the level of the inter-proximal papillae and the gingival margin (Fig. 3), the percentage of this area in relation to the total labial tooth surface could also be calculated if preferred (the shade of the red disc was not altered by the dampness from the small amount of saliva used for attachment).

For validation, comparison was made with an image analysis-based method by Ellis *et al.* (5) which had the benefit of being derived from an established method by Seymour *et al.* (6). This method involved taking photographic slides at a fixed focal distance and projecting them onto a screen which produced an enlarged image (×10) ready for scoring. The papilla on the labial aspects of the upper and lower anterior teeth were scored by Seymour's method attributing a score of 0–3 to each papilla, 0 = no gingival encroachment, 1 = mild encroachment and inter-dental papilla, 2 = moderate encroachment, involving lateral spread of papilla across the buccal tooth surface of less than one-quarter of the tooth width and a score of 3 = marked encroachment of papilla extending more than one-quarter of the tooth width. A total score was then calculated for each patient (max = 30 from the 10 papilla scored) and converted into a percentage overgrowth score.

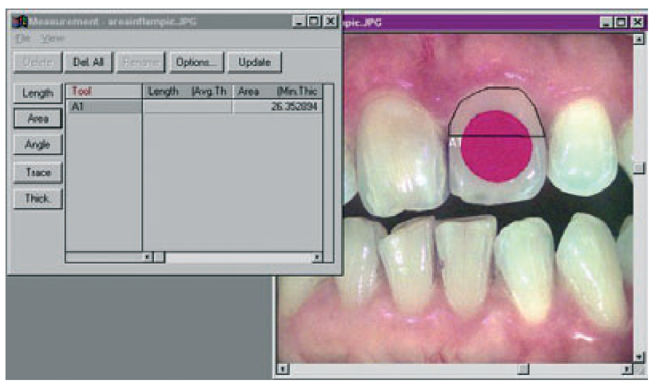

Fig. 3. Image of teeth showing portional tooth surface area measurement (mm^2) of an upper anterior tooth and the output menu after linear calibration.

Statistical methods

Descriptive statistics including the mean difference, SD and SE were calculated from the repeat measures of the proportion of tooth surface area described, red disc mean red pixel values and gingival mean red pixel values for both examiners. Bias was tested using a Student's *t*-test and illustrated using Bland Altman plots produced for all variables to test for unwanted size-difference relationships (9).

A two-tailed paired *t*-test was performed between first and second attendance data, with a significance level of 95% being selected (checked for normality). This was to detect any significant differences in gingival swelling.

The reliability of the method used to measure the relevant area and the mean red pixel value of the gingival and red calibration disc was calculated from duplicate images that were taken by each of the two examiners on the first clinical attendance of the patients (pretreatment). The subjects were removed and then repositioned in the apparatus between repeat images taken a few minutes apart so that total system error could be assessed. Repeat images had to be taken close together as minute changes may have occurred to the degree of overgrowth if a period of several hours had been left between repeat images. This would have negated the data for reliability purposes. Reliability was assessed according to the classification by Donner and Eliasziw (10) of the Fleiss' (11) coefficient of reliability (repeat measures taken during the patients' first attendance), which accounts for biological variation.

Comparison was made between the image analysis and the scoring system established by Ellis *et al.* (5) using Pearson's product–moment correlation coefficient.

Results

Table 1 presents the descriptive statistics for any measurement/operator bias for intra- and inter-operator reliability calculated to provide evidence for assessment of the reliability of the method including the average red calibration disc red pixel value, the average gingival redness pixel value and the portional tooth surface area. All *P*-values are greater than 0.05 inferring no significant difference between repeat measures.

Fleiss' coefficient of reliability data is shown in Table 2. All the values indicate high or excellent levels of method repeatability and reproducibility.

The summary descriptive statistics for the first and second attendances can be seen in Table 3. These data were used to evaluate any changes between the first and second attendance measurements using a two-tailed paired *t*-test. Table 4 shows that no significant differences were found.

Table 1. **Descriptive statistics for measurement of upper right and left central incisor areas for inflammation assessment and the average red RGB pixel level of the red discs and the visible gingivae (first attendance)**

	n	Mean difference	SD	SE	*t* (*P*)
Intra-operator repeatability					
O1UR	20	−0.013	1.880	0.427	0.977
O1UL	20	0.1740	1.460	0.331	0.605
O1 Disc	20	0.024	3.560	0.814	0.977
O1 GING	20	0.626	2.850	0.637	0.340
O2UR	20	−0.235	1.450	0.325	0.479
O2UL	20	−0.682	1.710	0.357	0.072
O2 Disc	20	−0.132	0.800	0.182	0.478
O2 GING	20	0.230	0.900	0.199	0.264
Inter-operator reproducibility					
IUR	20	−1.327	2.983	0.667	0.061
IUL	20	0.820	2.445	0.547	0.150
I Disc	20	0.010	6.606	1.477	0.995
I GING	20	4.759	12.812	2.865	0.059

O1 or O2 = operator 1 or operator 2 intra-comparison for the measurement of the portional tooth surface area. UR or UL = operator 1 or operator 2 intra-operator comparison for the upper right or upper left central incisors. Disc = operator 1 or operator 2 intra-operator comparison for the mean red pixel level of the red disc. GING = operator 1 or operator 2 intra-operator comparison for the mean red pixel level of the gingivae. IUR or IUL = inter-operator comparison for the measurement of the upper right or upper left central incisor for inflammation encroachment. I DOT = inter-operator comparison for the measurement of the mean red pixel level of the red dot. I GING = inter-operator comparison for the measurement of the mean red pixel level of the gingivae.

Table 2. **Fleiss' coefficient of reliability for measurement of upper right and left central incisor area for inflammation assessment and the average red RGB pixel level of a red disc and the visible gingivae**

	Intra-operator		
	Operator 1	Operator 2	Inter-operator
Area for inflammation study			
Upper right central incisors	0.975	0.989	0.947
Upper left central incisors	0.978	0.970	0.936
Average pixel red level			
Red disc	0.968	0.998	0.878
Gingival redness	0.987	0.998	0.769

R = 0.81–1.00 implies excellent reliability, *R* = 0.61–0.80 implies substantial reliability [Donner and Eliasziw (10)].

Pearson's correlation coefficient for comparison of the two gingival encroachment methods was −0.798 significant at the 0.01 level (two-tailed).

Discussion

The results in Table 1 show that no significant bias was observed for any of the variables. This is also seen both from

Table 3. **Summary data for first and second attendance for the assessment of gingival redness and red disc average red pixel level**

Variable	Occasion	*n*	Mean	±SD	SE
Gingival redness (red disc) (RGB mean red pixel value)	First visit	20	−16.775	17.916	4.006
	Second visit	20	−15.521	18.061	4.039
Upper right central incisor area (mm^2)	First visit	20	21.759	9.900	2.214
	Second visit	20	20.982	10.109	2.26
Upper left central incisor area (mm^2)	First visit	20	22.952	8.120	1.816
	Second visit	20	22.812	8.742	1.955

Table 4. **Two-tailed paired t-test results for changes of gingival redness average red pixel level and tooth area before and after treatment**

Variable	*n*	Mean	±SD	SE	*t*-Test (*P*)
Gingival redness (red disc value) (RGB mean red pixel value)	20	−1.254	22.92	5.12	0.809
Upper right central incisor area surrounded by gingiva) (mm^2)	20	0.777	4.717	1.055	0.470
Upper left central incisor area (surrounded by gingiva) (mm^2)	20	0.141	4.565	1.021	0.892

the reliability *t*-test values ($P > 0.05$ to show no significant difference or bias between each measure) and that all the mean differences are less than 1.96 times the standard error. The Bland Altman plots produced no indicative regular data patterns or trends suggesting size/error relationships.

The reliability study showed that the red calibration disc and gingival redness values were the most reliable with excellent Fleiss coefficient of reliability *R* values of 0.968–0.998 (Table 2).

Subtraction of the red disc mean red pixel value from that of the visible gingival on all visits enabled any systematic differences to be accounted for. Therefore, the remaining change in gingival value could be attributed to changes in inflammation.

There was no evidence of a statistically significant difference at the 95% level between first and second visits (Table 4). A few of the cases in this study when seen after a year, showed slight gingival redness level changes and reduced coverage of tooth surfaces from the measurement of the tooth area. However, most cases showed little change in short-term assessments. The lack of change in the gingival colour and area of a patient may be accounted for by failure to achieve effective professional root surface cleaning, poor patient compliance, medical conditions affecting the host response or medication having an adverse effect on gingival inflammation and encroachment. However, in the group of patients examined, whilst good measurable improvements were seen for some patients, the overall results within the group may have been adversely affected by a number of patients having a poor response to treatment.

One individual typical of those investigated is presented in Fig. 4. The gingival inflammation redness (average red pixel value) rose by 11.9, as the inflammation declined indicated by an increase in RGB value (as inflammation reduces the red level darkens giving an increase in the RGB average red pixel level and vice versa). The upper right central incisor, however, showed a decrease in area between the gingival margin and inter-proximal papilla perimeter of 3.98 mm^2 signifying inflammatory enlargement compared with the upper left central incisor that showed an increase in area of 0.36 mm^2 signifying a reduced inflammatory response. Overall, the inflammation level appears to have dropped with a localized increased inflammatory response around the upper right central incisor.

Comparison of our tooth area approach for assessing gingival encroachment and the Ellis *et al.* (5) established scoring technique showed a strong correlation (−0.798), validating this part of the technique. The image analysis system, however, by design should be able to detect smaller changes in gingival encroachment levels due to its objective non-index system approach.

The variables chosen for this study were derived with the current technology in mind and the need to provide objective

Fig. 4. A patient, showing gingival colour change, imaged on two occasions with a month between visits. The left image shows the pretreatment condition and the right image shows the post-treatment presentation.

results as far as possible. Therefore, the choices of an assessment of colour and area change seemed reasonable. The standard disc was used as there are minute positional and illumination differences between images that cannot be totally removed even with our standardized system.

The authors accept that due to imaging restrictions there may be some subjectivity in the redness assessment due to selection of the area of interest, although the reliability results proved sound. The authors also accept that this method can only assess the anterior part of the oral cavity and often there is severe inflammation in the posterior part.

Although the main inflammatory response is pronounced at and largely limited to the gingival margin, the method used the red pixel values from the total gingival area. This was to include any subtle changes from other regions of gingiva. As the method was reliable, this inclusion should not mask the changes at the gingival margin.

From a practical point of view, this system could be used to give patients a motivating pictorial indication of their changing periodontal/gingival state as well as providing objective data to match those changes. Similarly, permanent database of images is created that are useful for objectively assessing reliability, reviewing extraneous results or for further research studies. Of advantage, the method does not require qualified clinicians unlike many of the existing indices. Nevertheless, image analysis is limited to recording of colour change and area only and as such traditional clinical assessments and examinations would still need to be undertaken.

Conclusions

The method is highly reliable, both for measurement of the tooth area within the perimeter of the inter-proximal papillae and the gingival margin and assessing the average RGB red pixel value. The method has proved easy to perform after initial training and provides an additional tool for assessing gingival health individually in a patient or in a clinical trial.

References

1 Schour I, Massler M. Gingival disease in post-war Italy (1945). I. Prevalence of gingivitis in various age groups. *J Am Dent Assoc* 1947; **35:** 475–482.

2 Parfitt GJ. A five year longitudinal study of the gingival condition of a group of children in England. *J Periodontol* 1957; **26:** 26–32.

3 Russell AL. A system of classification and scoring for prevalence surveys of periodontal disease. *J Dent Res* 1956; **35:** 350–359.

4 Harris TH, Ewalt JR. Complications following the use of sodium diphenyldantoinate (dilantin) therapy. *J Okla State Med Assoc* 1942; **35:** 365–370.

5 Ellis JS, Seymour RA, Robertson P, Butler TJ, Thomason JM. Photographic scoring of gingival overgrowth. *J Clin Periodontol* 2001; **28:** 81–85.

6 Seymour RA, Smith DG, Turnbull DN. The effects of phenytoin and sodium valproate on the periodontal health of adult epileptic patients. *J Clin Periodontol* 1985; **12:** 413–419.

7 Smith RN, Brook AH, Elcock C. The quantification of dental plaque using an image analysis system: reliability and validation. *J Clin Periodontol* 2001; **28:** 1158–1162.

8 Smith RN, Rawlinson A, Lath D, Elcock C, Walsh TF, Brook AH. The quantification of dental plaque on lingual tooth surfaces using image analysis: reliability and validation. *J Clin Periodontol* 2004; **31:** 569–573.

9 Altman DG. *Some Common Problems in Medical Research, Practical Statistics for Medical Research*, London, Chapman & Hall, 1996.

10 Donner A, Eliasziw M. Sample size requirements for reliability studies. *Stat Med* 1987; **6:** 441–448.

11 Fleiss JL. *Reliability of Measurement, the Design and Analysis of Clinical Experiments*. London, Wiley and Sons, Inc., 1986.

Copyright of International Journal of Dental Hygiene is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.