J Clin Periodontol 2007; 34: 969–976 doi: 10.1111/j.1600-051X.2007.01137.x

Journal of Clinical Periodontology

Root coverage assessment: validity and reproducibility of an image analysis system

Stéphane Kerner[1], Daniel Etienne[1], Jacques Malet[1], Francis Mora[1], Virginie Monnet-Corti[2] and Philippe Bouchard[1]

[1]Department of Periodontology, Service of Odontology, Hôtel-Dieu Hospital, AP-HP, Paris 7 – Denis Diderot University, U.F.R. of Odontology, Paris, France; [2]Department of Periodontology, Faculty of Odontology, University of Marseille, Marseille, France

Kerner S, Etienne D, Malet J, Mora F, Monnet-Corti V, Bouchard P. Root coverage assessment: validity and reproducibility of an image analysis system. J Clin Periodontol 2007; 34: 969–976. doi: 10.1111/j.1600-051X.2007.01137.x.

Abstract

Aim: The aim of this methodological study was to validate a new method for root coverage evaluation following periodontal plastic surgery.

Material and Methods: Thirty recessions were treated in 21 consecutive patients, using a subepithelial connective tissue graft technique. Clinical measurements and photographs were taken at baseline and 12 ± 6 months after treatment. The mean percentage of root coverage for linear and surface area measurements was calculated using conventional clinical evaluation, and compared with ImageJ, a public domain Java image processing program. Bland–Altman plots were used for assessing repeatability and agreement between clinical and ImageJ measurements. The strength of the relationship was calculated using the Pearson product moment correlation coefficient.

Results: The repeatability of ImageJ was excellent for both linear and surface area measurements. The agreement between clinical and ImageJ measurements was good for the linear evaluation, showing lower and upper limits of -13.14% and 17.42%, respectively. Significant correlations ($p < 0.001$) were found between clinical and ImageJ measurements, ranging from 0.93 to 0.94 for linear evaluation, and from 0.89 to 0.90 for surface evaluation.

Conclusions: The outcomes of this study show that the ImageJ analysis is a reliable, reproducible method to evaluate the percentage of root coverage after periodontal plastic surgery, when a midfacial linear measurement is used.

Key words: gingival recession/therapy; image analysis; measurement reproducibility; methodology; validity

Accepted for publication 20 July 2007

The efficacy of the treatment of recession defects is achieved by the quantification of the amount of root coverage after periodontal plastic surgery. Clinical trials normally record the following outcome measurements: (1) change in recession depth (RD), (2) change in clinical attachment level (CAL), (3) mean percentage of root coverage, (4) frequency of complete root coverage (Roccuzzo et al. 2002). Consequently, root coverage assessment may be evaluated through absolute values and/or proportions. RD is measured, with a conventional manual periodontal probe, as the distance from the cemento-enamel junction (CEJ) to the gingival margin (GM). This measurement is subject to two types of potential systematic errors: (1) the accuracy of periodontal probing (Hefti 1997), and (2) the localization of the anatomic CEJ (Zucchelli et al. 2006). In order to reduce these potential errors, the examiners involved in clinical studies need to be calibrated for probing measurements before and sometimes – but rarely – during the study. Despite these limitations inherent to probing measurement, clinical measurements are still used as a ''gold standard'' to evaluate the mean percentage of root coverage in clinical trials dealing with the evaluation of gingival recession defects.

Thus, it is of interest to develop new methods that may simplify the experimental protocols and may potentially reduce the measurement errors. Recent developments in image analysis allow image quantification, and are appropriated to further explore visual evaluation of the gingival recession defect model in terms of percentage of root coverage after a surgical treatment of the defect.

It has been also suggested that RD evaluation based on a single vertical

Conflict of interest and source of funding statement

The authors declare that they have no conflict of interests.

No external funding, apart from the support of the authors' institution, was available for this study.

linear midfacial measurement may be a less-accurate criterion in the evaluation of the mean percentage of root coverage than the measurement of the recession surface area (RSA) (Bouchard et al. 1997, Rosetti et al. 2000).

The aim of this study was (1) to validate a new measurement technique of the mean percentage of root coverage (RD) by using image analysis compared with the conventional clinical evaluation, and (2) to evaluate the use of RSA measurement with this methodology.

Material and Methods

Study sample and experimental design

Subjects were selected from regular routine patients attending the Service d'Odontologie Hôtel-Dieu, and referred for root coverage procedure at the Department of Periodontology, Paris 7 – Denis Diderot University, France. The study was conducted in accordance with the Helsinki Declaration of 1975, as revised in 2000. No Ethical Committee approval was required because all the patients included in the study did not undergo procedures other than those normally required by the guidelines of the department. Thus, the study protocol involved a standard initial therapy to establish optimal plaque control [dichotomous plaque index (PI) $\leqslant$ 20%] and gingival health conditions (BOP $\leqslant$20%), followed by root coverage therapy, post-operative plaque control, and final evaluation, at least 6 months after the surgical procedure. The policy of the department for root coverage procedures calls for clinical measurements of the recession defect and a photograph of the defect the day of the surgery and another at least 6 months later for control. Patients had to be free of gingivitis or periodontitis. Patients were aware that data from their examination would be used in a research study and that no person would be identified in any publication. Only patients who agreed that their data could be used were included in the analysis.

Patients chosen for the validation methodology had to be treated with a subepithelial graft procedure, and had to meet the following inclusion criteria:

- Males and females must be at least 18 years of age.
- The patients must be in good general health without any systemic diseases.
- The patients must have at least one buccal class 1 or 2 Miller's gingival recession defect.
- The recession defects must be limited to maxillary and mandibular premolars, canines, and incisors.
- The experimental teeth must be free of endodontic lesions and of caries or restorative dentistry in the defect area at the date of the surgical procedure and at the date of the control.
- The experimental teeth must be documented with high-quality photographs at the date of the surgery and at least 6 months-later for the control.

The following criteria excluded patients from the study:

- Any medical conditions that could interfere with normal healing, including current pregnancy at the time of the surgical procedure.
- Palatal and lingual gingival recessions.
- Molar teeth.
- Teeth with untreated endodontic or cariologic lesions.
- Lack of visibility of the CEJ.

Sample size calculation was performed using a statistical software program (PASS 2005, NCSS Statistical Software, Kaysville, Utah). Using the recession as the statistical unit, a sample size of 17 was calculated to achieve 90% power at the two-sided 5% level to detect a difference of 0.70 between the null hypothesis correlation and the alternative hypothesis correlation. To strengthen the results of the study, we decided to set the sample size at 30. Thus, 30 recessions were treated from October 1999 to June 2002 in 21 consecutive patients meeting the selection criteria, three males and 18 females, mean age 38.7 years ($\pm$ 8.5 years; age range, 21–61 years). The mean follow-up was 1 year $\pm$ 6 months. Ten Miller's class I and 20 Miller's class II recessions were included in the study. Twenty-three maxillary and seven mandibular teeth including four incisors, 12 canines, and 14 pre-molars were subjected to analysis.

Surgical procedure

All surgical procedures were performed by the same operator (P. B.). Details of the surgical procedure have been described previously (Bouchard et al. 1997). Briefly, following intra-crevicular incisions through the bottom of the crevice, a partial thickness flap was raised at the buccal aspects of the tooth. Two mesial and distal vertical incisions were made external to the interdental papillae. The flap was extended coronally until it covered the entire area of recession up to the CEJ. A wedge of connective tissue was harvested from the palate and its small band of epithelium excised. The connective tissue graft was placed on the denuded root and the flap coronally positioned as high as possible in order to completely immerse the graft. The graft was secured with interrupted sutures. Routine post-operative care included analgesics, lack of brushing on the surgical area until suture removal (14 days), and the use of chlorhexidine rinse twice daily for 2 weeks.

Clinical measurements

The clinical measurements were recorded at baseline and at 6 months to 2 years post-operatively by the operator (P. B.). A straight manual probe graduated in 1 mm increments was used (PCP UNC-15, Hu-Fredy®, Chicago, IL, USA).

Intra-examiner reliability is performed annually by a calibration session. The training exercise consists of the examination of two practice calibration subjects (half-mouth). Repeated measurements are made on each calibration subject, following at least a 10-min. time lapse, until the agreement within 1.0 mm between two passes is more than 95%. In the present study, a reading that fell between 2 mm increments was rounded down.

The RD was recorded in the mid-buccal part of the tooth, from the CEJ to the GM. The probing depth (PD) was measured from the GM to the tip of the probe.

Image measurements

Acquisition and storage of recession images

Photographs (magnification ratio: 1/1) of the recessions were made pre-operatively and again at least 6 months later (Nikon Medical 120, Tokyo, Japan) by the operator. No custom-built frame or special equipment was used. Views were taken according to the standard shooting protocol used in the

department, which ensures that all photographs are made in the axis to facilitate proper recording with minimal distortion. When required, a glass mirror was used at baseline and control.

Slides were digitalized under 300 dpi with a scanner, and displayed using Adobe® Photoshop® software (version 7.0, Adobe Systems Europe Ltd., Uxbridge, UK) by the examiner (S. K.). Two cases were excluded from the validation study because the CEJ was not completely detectable and/or the pre-operative and post-operative views had distortion that impeded comparison. Cases were stored until 30 recessions were properly selected, corresponding to 60 photographs (baseline+follow-up).

ImageJ analysis

RD and RSA were analysed using ImageJ for windows. ImageJ is a public domain Java image processing program, which calculated area and pixel value statistics for user-defined selections (Abramoff et al. 2004).

The tooth of interest was cropped, and the following two lines were drawn on a graphic tablet: (1) a straight apico-coronal vertical line from the most apical point of the RD to the most coronal portion of the crown edge; (2) a mesio-distal horizontal line, at the widest part of the crown. These two lines were used as references to check the reproducibility of the magnification. Then, the CEJ and the contour of the recession were drawn (Fig. 1A).

To ensure the examiner's blinding, photographs were coded independently into a spreadsheet using a proprietary randomization programme (under Excel, Microsoft, Redmond, WA). Unblinding was performed after calculation completion for merging with clinical worksheets.

RD was calculated using the apico-coronal line as a reference. RSA was calculated as the area within the contour of the denuded root (Fig. 1B). Duplicate measurements were made for both RD and RSA. Results were given in pixel value.

Percent root coverage was calculated for RD and for RSA according to the following standard formulae:

- [(pre-operative RD – post-operative RD)/(pre-operative RD)] × 100,
- [(pre-operative RSA – post-operative RSA)/(pre-operative RSA)] × 100.

Intra-examiner calibration

Intra-examiner variability was evaluated by triplicate measurements of three recession defects. The aim of this calibration training was to detect a maximum of 5% of the coefficient of variation for both root surface recession and linear evaluation. Upon completion of the training, the coefficients of variation for root surface evaluation and linear evaluation were 0.98%, 1.31%, 1.82%, and 1.18%, 1.19%, 2.05%, respectively.

Data analysis

Data collected were organized into a spreadsheet using a computer program (Excel, Microsoft, Redmond, WA). Bland–Altman plots were used for assessing repeatability of the image analysis as well as the agreement between the clinical measurements and image measurements (Bland & Altman 1986). The limits of agreement were calculated using the following formula: $\bar{d} \pm 2s$, where $\bar{d}$ is the mean difference and s is the standard deviation of the differences.

The strength of the relationship between linear and surface measurements was assessed by using the Pearson product moment correlation coefficient.

Statistical analysis and data management were performed using a statistical software program (NCSS 2004, NCSS

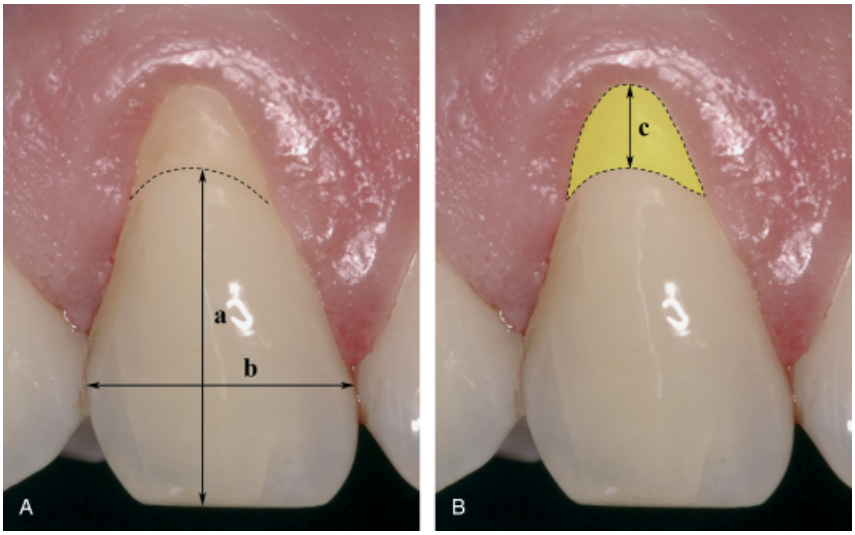

Fig. 1. Measurements with ImageJ. (A) Vertical (a) and horizontal (b) lines are drawn to check the reproducibility of the pictures. (B) The RD (line c) and RSA (yellow surface) allowed for evaluating recession defects.

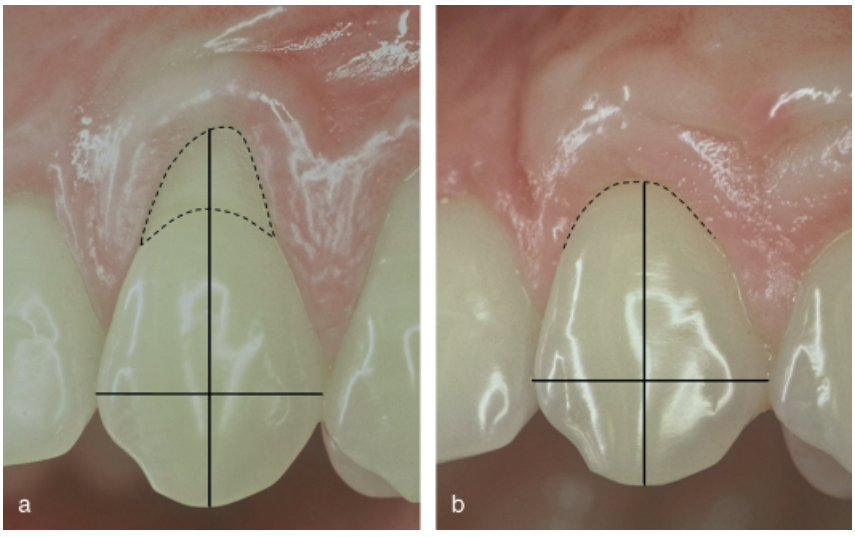

Fig. 2. Preoperative (a) and postoperative (b) views of a clinical case with complete root coverage.

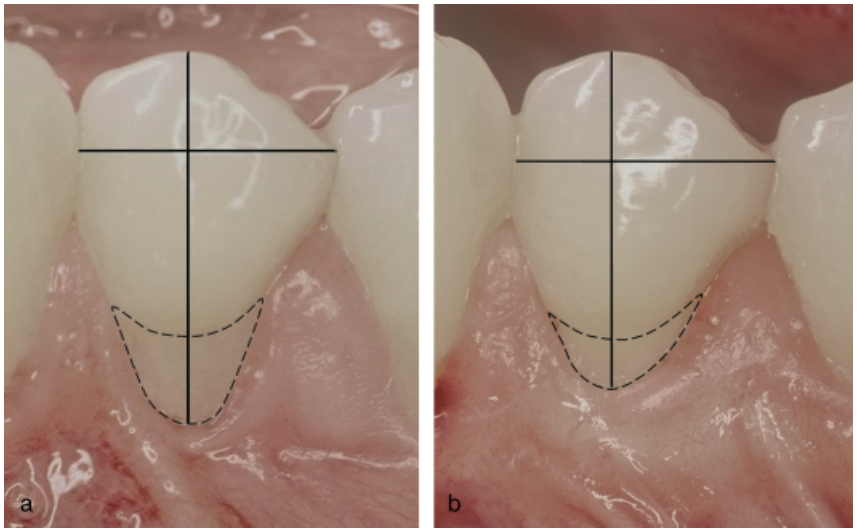

Fig. 3. Preoperative (a) and postoperative (b) views of a clinical case with partial root coverage. The clinical value is 40%. Using ImageJ, the values of RD is 42% and RSA is 41%.

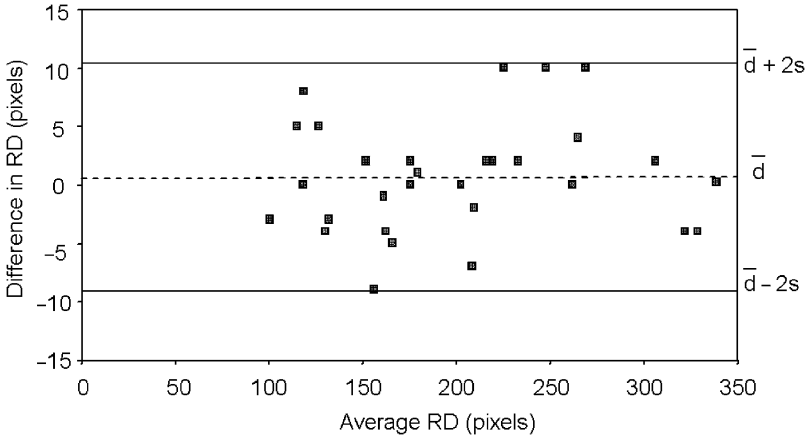

Fig. 4. Bland–Altman plot of repeated measures of RD using ImageJ. The mean difference is 0.63 pixels and the limits of agreement are − 9.09 and 10.35.

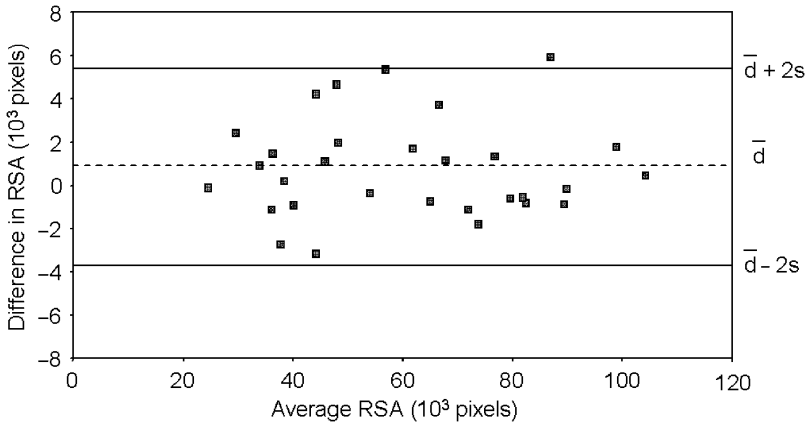

Fig. 5. Bland–Altman plot of repeated measures of RSA using ImageJ. The mean difference is 726.8 pixels and the limits of agreement are − 3817.58 and 5271.18.

Statistical Software). The level of statistical significance was set at 0.05. Means and standard deviations were determined to describe the data.

Results

The mean clinical root coverage from baseline to control results was 73.83 ± 21.33%. Complete root coverage was achieved in nine teeth of the 30 defects. The mean clinical RD decreased from 4.10 ± 0.92 to 1.10 ± 0.96 mm.

The mean root coverage for RD using the Image analysis was 71.62 ± 21.23% and 71.69 ± 21.26% for the first and the second measurements, respectively. Using the RSA variable, the mean root coverage was 68.14 ± 21.81% and 68.16 ± 22.53% for the first and the second measurements, respectively (Figs 2 and 3).

Repeatability of the ImageJ method

The mean RD for duplicate measurements using ImageJ was 200.88 ± 68.17 pixels. Figure 4 indicates the repeatability for repeated measures of RD. The lower and upper limits of agreement were − 9.09 and 10.35 pixels, respectively. The mean RSA was 60,660.93 ± 22,275.58 pixels. Figure 5 shows the repeatability for repeated measured of RSA using ImageJ. The lower and upper limits of agreement were − 3817.58 and 5271.18 pixels, respectively. These data indicate an excellent repeatability of the method for the two variables RD and RSA.

Agreement between clinical and ImageJ measurements

Figure 6 indicates the agreement between clinical measurements and ImageJ measurements for RD. The lower and upper limits of agreement were − 13.14% and 17.42%, respectively, showing a good agreement between the two methods. Figure 7 shows the agreement for RSA measurements. The lower and upper limits of agreement were − 13.80% and 25.14%, respectively. These data show a moderate agreement between clinical and ImageJ for RSA measurements.

Strength of the relation between clinical and ImageJ measurements

Figures 8 and 9 show the linear relationship between the mean percentage of clinical root coverage and the mean percentage of RD reduction measured with ImageJ. The correlations were significant ($p < 0.001$) and ranged from 0.93 for the first measurement to 0.94 for the second measurement. Figures 10 and 11 show the relationship between the percentage of clinical root coverage and the mean percentage of RSA

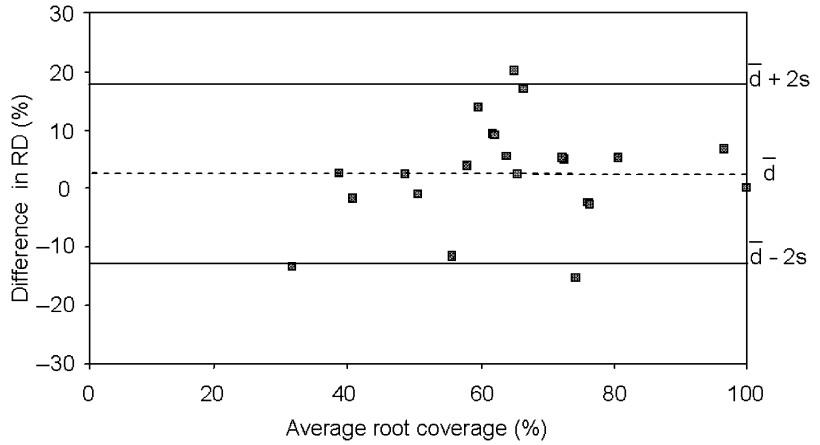

Fig. 6. Bland–Altman plot of difference against RD mean for root coverage data by clinical measurements and ImageJ. The mean difference is 2.14% and the limits of agreement are − 13.14 and 17.42%.

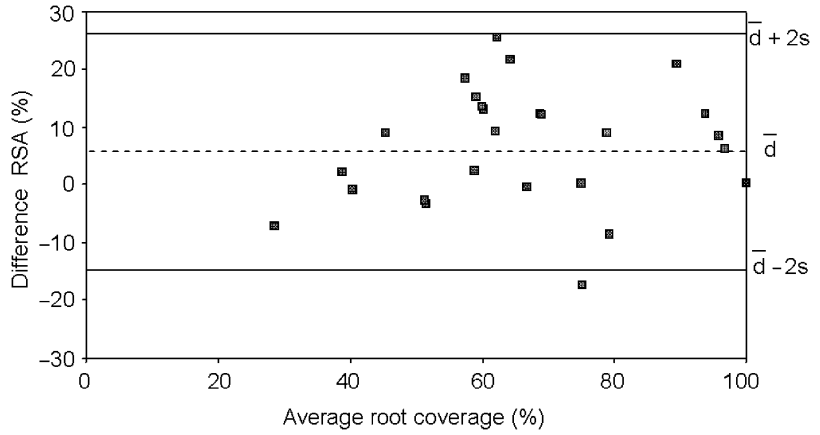

Fig. 7. Bland–Altman plot of difference against RSA mean for root coverage data by clinical measurements and ImageJ. The mean difference is 5.67% and the limits of agreement are − 13.80 and 25.14%.

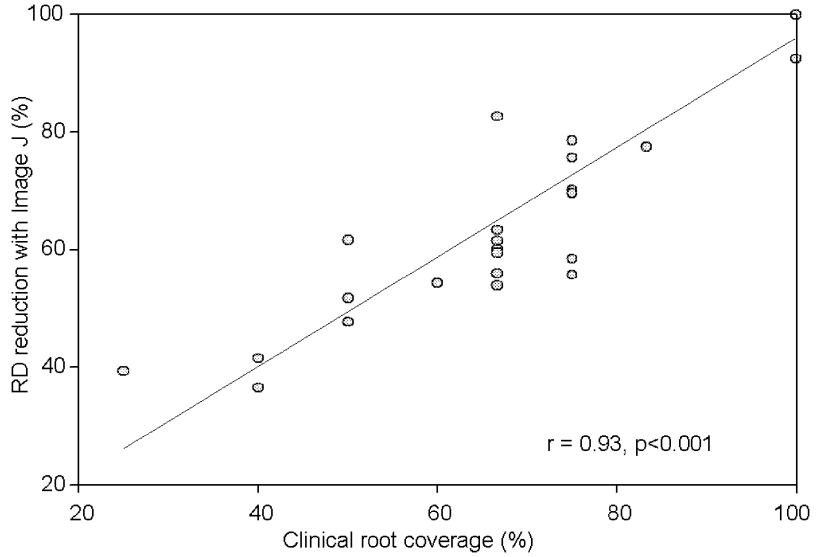

Fig. 8. Linear regression plot (first measurement) between the mean percentage of clinical root coverage and the mean percentage of RD reduction measured with ImageJ for the first measurement. Square root of the MSE = 0.077; Estimated slope = 0.93 $CI_{95\%}$ [0.79–1.07]; Estimated Intercept = 0.029 $CI_{95\%}$ [− 0.076–0.134].

reduction. Again, the correlation was highly significant ($p < 0.001$) with correlation coefficients ranging from 0.89 for the first measurement, to 0.90 for the second measurement. These results indicate that RD and RSA measured with ImageJ are highly correlated with clinical measurements.

Discussion

The outcomes show that the ImageJ analysis is a reliable, reproducible method to evaluate the amount of root coverage after periodontal plastic surgery.

Compared with the gold standard, i.e. clinical measurement of the recession, it may be assumed that this new method is more convenient in daily practice and less subjective in clinical research. ImageJ is a public domain Java image processing program, based on NIH Image, which can be downloaded on a website (http://rsb.info.nih.gov/ij/download.html) (Rasband 1997–2006). It may be used by any practitioner trained on the program.

The major advantage of this method is the use of information, which is documented in any case. As seen in the ''Materials and methods'' of the present study, the reproducibility does not imply the use of additional support other than that which is normally used in a routine approach of these treatments. In routine practice, the practitioners are not calibrated. Thus, clinical measurements are subjected to intra-examiner error. The present method, when inclusion/exclusion criteria are respected, is not subject to this major bias and allows the use of these measurements in the personal database of the practitioner, to evaluate (1) his or her individual performance over time, and (2) patients' follow-up. On a research basis, this method cannot replace the traditional and readily available clinical measurements in randomized clinical trials, because it does not take into account the absolute value of the recession. However, from the perspective of multicentre retrospective studies aiming to evaluate the outcomes of daily use of different techniques in a routine approach, this new method may be useful because the inter-examiner calibration is not mandatory. Thus, individual databases may be compiled, allowing for the analysis of databases larger than those used for randomized clinical trials.

Digital measurements have been used in root coverage clinical studies (Rosetti

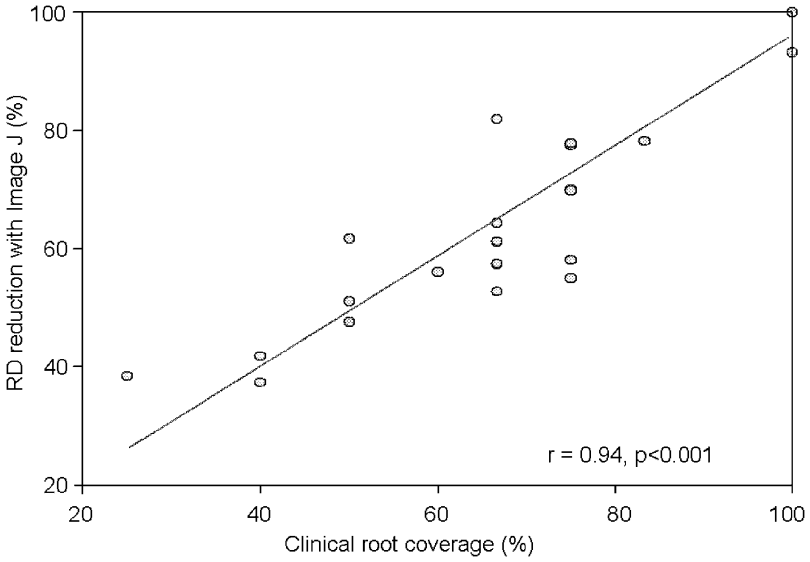

Fig. 9. Linear regression plot (second measurement) between the mean percentage of clinical root coverage and the mean percentage of RD reduction measured with ImageJ for the first measurement. Square root of the MSE = 0.077; Estimated slope = 0.93 $CI_{95\%}$ [0.80–1.07]; Estimated Intercept = 0.029 $CI_{95\%}$ [− 0.076–0.133].

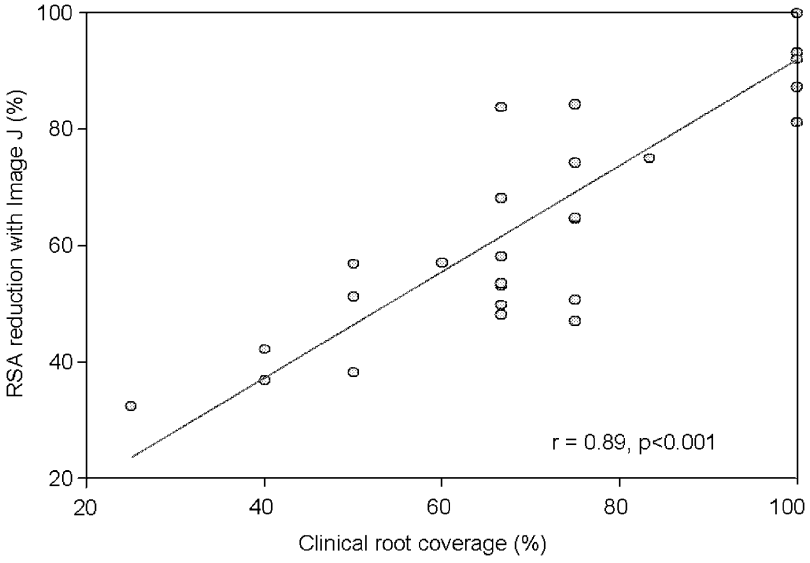

Fig. 10. Linear regression plot (first measurement) between the mean percentage of clinical root coverage and the mean percentage of RSA reduction measured with ImageJ for the first measurement. Square root of MSE = 0.10; Estimated slope = 0.91 $CI_{95\%}$ [0.73–1.09]; Estimated Intercept = 0.01 $CI_{95\%}$ [− 0.13–0.15].

et al. 2000, Saletta et al. 2001, Yotnuengnit et al. 2004). However, none of these studies have evaluated the reproducibility of the method. These studies used a specific designed protocol to compare clinical and digital measurements, making its use difficult in routine practice. In this study, no purpose-designed system using a sophisticated custom-built frame was used. Photographs of the recessions were made in normal clinical situations. Thus, the system does not require additional clinician time and/or volunteer time.

The photographs are part of the standard clinical requirements when dealing with aesthetics. Today, no database of images exists, which can be later re-evaluated or used for further research. Most of the clinical trials include a limited number of recession/subjects due to protocol requirements. It would be of interest to investigate the outcomes of various root coverage techniques performed in standard conditions (chair side), including a large number of data. It must be kept in mind that the aim of clinical trials is to evaluate the clinical effect of a specific variable on a designed procedure. In daily practice, it is hard to conclude from these trials that one procedure is better than another. Furthermore, most of these studies cannot lead to significant conclusions because of small sample size.

In the present study, clinical measurements and surgical procedure were performed by the same operator. This fact does not affect the evaluation of reproducibility, which is the aim of the study, but the overall clinical results must be viewed in this perspective.

It may be assumed that the risk of error with ImageJ is less important than with the standard clinical measure. A typical value of the clinical RD is 4 mm. This value is rounded to the nearest mm. This corresponds to a potential 25% error of variation of the measurements. With ImageJ, the corresponding potential error for RD = 4 mm is 1/330 pixels, corresponding to 0.003% (Fig. 12). It may be considered that clinical measurements are less precise than the ImageJ measurements. One can assume that the use of a custom stent to perform clinical measurements may improve the accuracy of the results. However, very few root coverage studies use this device for clinical parameters assessment. The reason that may be invoked is the lack of evidence in the literature showing an advantage in the use of acrylic stents compared with standard clinical measurements in the specific evaluation of RD.

We found moderate agreement between clinical measurements and ImageJ for RSA measurements. In this study, the view angle was not controlled by a specific device. Surface measurement is more sensitive to the distortion than a linear measurement, and may explain the difference in agreement between RD and RSA. It would be of interest to evaluate the root surface modification according to the variation of the angle of the snapshot. However, this result indicates that the present methodology may preclude the use of RSA measurement to evaluate root coverage with image capture without the use of a custom-built frame.

The major limitation of the ImageJ evaluation is that the absolute value of

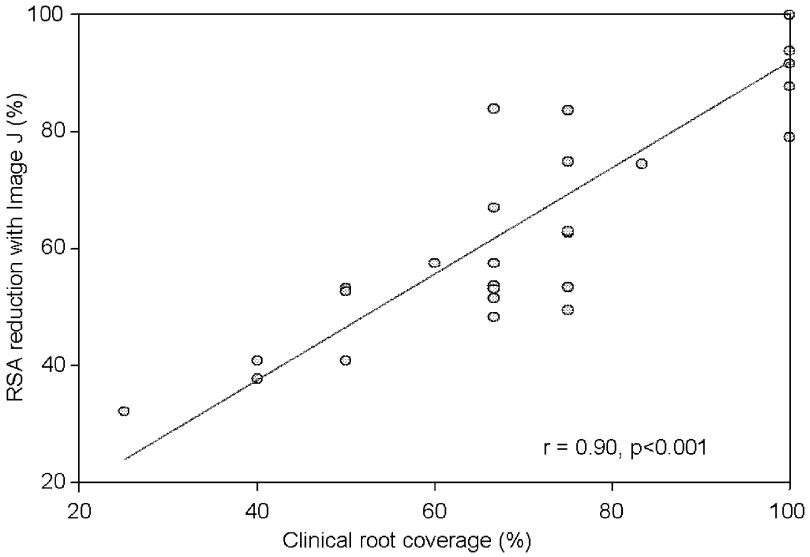

Fig. 11. Linear regression plot (second measurement) between the mean percentage of clinical root coverage and the mean percentage of RSA reduction measured with ImageJ for the first measurement. Square root of the MSE = 0.10; Estimated slope = 0.91 $CI_{95\%}$ [0.73–1.08]; Estimated Intercept = 0.01 $CI_{95\%}$ [− 0.12–0.15].

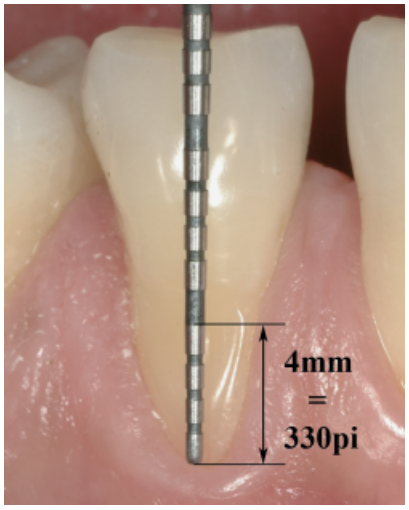

Fig. 12. Clinical and ImageJ evaluation of a 4 mm recession depth.

the RD cannot be recorded. Thus, it is not possible with this technique to classify the defects according to the magnitude of the recession. However, the landmark report of the 3rd European Workshop does not indicate the value of the recessions at baseline, but the percentage of root coverage and the mean difference between baseline and postop (Roccuzzo et al. 2002). Futhermore, most of the studies indicate a range of RD at baseline of 2.2–5.5 mm. Miller's classification seems to be the best indicator for root coverage limitations (Miller 1985).

Another limitation of the method is the quality of the photographs. Poor-quality photographs cannot be analysed due to the difficulty in defining the CEJ location. It may be assumed that the use of a digital camera would improve the image analysis and reproducibility.

In conclusion, the ImageJ analysis provides a simple and reliable method of quantifying root coverage without the need for complex device. It is a useful, fast, sensitive technique, and can be advised for clinicians and researchers for the evaluation of the percentage of root coverage. It is objective and may provide an image database for future research. Today, its use should be limited to RD evaluation. Further research is needed to evaluate the influence of image distortion on RSA measurements.

Acknowledgements

The authors wish to thank Dr. Pierre Boutouyrie, Department of Pharmacology, Hôpital Européen Georges Pompidou, INSERM U 652, AP-HP, France, for his interest and assistance in this study.

References

Abramoff, M. D., Magelhaes, P. J. & Ram, S. J. (2004) Image processing with ImageJ. *Biophotonics International* **11**, 36–42.

Bland, J. M. & Altman, D. G. (1986) Statistical methods for assessing agreement between measurement. *Lancet* **1**, 307–310.

Bouchard, P., Nilveus, R. & Etienne, D. (1997) Clinical evaluation of tetracycline HCl conditioning in the treatment of gingival recessions. A comparative study. *Journal of Periodontology* **68**, 262–269.

Hefti, A. F. (1997) Periodontal probing. *Critical Reviews in Oral Biology and Medicine* **8**, 336–356.

Miller, P. D. (1985) A classification of marginal tissue recession. *The International Journal of Periodontics and Restorative Dentistry* **5**, 8–13.

Rasband, W. S. 1997-2006 *ImageJ*. Bethesda, Maryland, USA: US National Institutes of Health, http://rsb.info.nih.gov/ij/ [accessed October 30, 2006].

Roccuzzo, M., Bunino, M., Needleman., I. & Sanz, M. (2002) Periodontal plastic surgery for treatment of localized gingival recessions: a systematic review. *Journal of Clinical Periodontology* **29** (Suppl. 3), 178–194.

Rosetti, E. P., Marcantonio, R. A., Rossa, C., Chaves, E. S., Goissis, G. & Marcantonio, E. (2000) Treatment of gingival recession: comparative study between subepithelial connective tissue graft and guided tissue regeneration. *Journal of Periodontology* **71**, 1441–1447.

Saletta, D., Pini Prato, G., Pagliaro, U., Baldi, C., Mauri, M. & Nieri, M. (2001) Coronally advanced flap procedure: is the interdental papilla a prognostic factor for root coverage? *Journal of Periodontology* **72**, 760–766.

Yotnuengnit, P., Promsudthi, A., Teparat, T., Laohapand, P. & Yuwaprecha, W. (2004) Relative connective tissue graft size affects root coverage treatment outcome in the envelope procedure. *Journal of Periodontology* **75**, 886–892.

Zucchelli, G., Testori, T. & De Sanctis, M. (2006) Clinical and anatomical factors limiting treatment outcomes of gingival recession: a new method to predetermine the line of root coverage. *Journal of Periodontology* **77**, 714–721.

Address:
Pr. Philippe Bouchard
U.F.R. d'Odontologie
5, rue Garancière
75006 Paris
France
E-mail: phbouch@noos.fr

Clinical Relevance

Scientific rationale for the study: Prospective clinical trials dealing with root coverage therapies are subjected to low sample size. Therefore, it would be of interest to develop new evaluation methodologies allowing larger sample size. In the present study, we used ImageJ, a free Java image processing program, to analyse pre-operative and post-operative photographs of treated recession defects.

Principal findings: The results of our study show that the ImageJ analysis is a reliable, reproducible method to evaluate the percentage of root coverage after root coverage therapy.

Practical implications: ImageJ analysis is a fast and sensitive method that can be advised for clinicians and researchers in root coverage evaluation. This validation should allow for the development of image databases, which could employ larger sample sizes in data analysis.