

Assessment of Intra- and Inter-examiner Reproducibility of Probing Depth Measurements with a Manual Periodontal Probe

Ardeshir Lafzi^{1*} • Adileh Shir Mohammadi² • Amir Eskandari² • Sohrab Pourkhamneh³

1. Associate Professor, Department of Periodontics, Faculty of Dentistry, Tabriz University of Medical Sciences, Iran.

2. Assistant Professor, Department of Periodontics, Faculty of Dentistry, Tabriz University of Medical Sciences, Iran.

3. Post-graduate Student, Department of Orthodontics, Faculty of Dentistry, Tabriz University of Medical Sciences, Iran.

* Corresponding author: E-mail: lafzia@tbzmed.ac.ir

Abstract

Background and aims. The main purpose of this study was to assess intra- and inter-examiner reproducibility of probing depth measurements with a manual periodontal probe.

Materials and methods. In this study, 32 dental students in Tabriz Faculty of Dentistry with normal periodontium were evaluated. Each tooth of the upper right quadrants, except the third molars, was examined. Probing depths were measured in 6 surfaces of each tooth (mesiofacial, midfacial, distofacial, mesiolingual, midlingual and distolingual). Each patient was examined by two examiners (two periodontists) in two sessions with an interval of 7-10 days. A total of 218 teeth and 1295 surfaces were examined.

Results. Intra-examiner measurements showed no statistically significant differences, while the differences of inter-examiner measurements were statistically significant. Intra-examiner reproducibility was more than the inter-examiner one. Measurements for anterior region, facial and mid-facial/mid-lingual surfaces were more reproducible than posterior, lingual and proximal surfaces.

Conclusion. Probing depth measurements with a conventional probe have an appropriate reproducibility in clinical settings, although variations between examiners may affect the reproducibility, especially when great accuracy is required.

Key words: Clinical probing depth, examiner, periodontal probe, reproducibility.

Introduction

Probing depth measurement has an important diagnostic role in the assessment of presence, severity and progression of periodontal disease and its treatment. Therefore, accuracy in probing depth measurement is very important.¹ Periodontal probes are the most widely accepted instruments for evaluating the

gingival status and periodontal health of dental patients, enabling the clinician to determine probing depth.¹ Several factors influence the reproducibility of probing depth measurements, which can potentially result in error. Inflammatory state of the gingiva, force, angulation, position and the diameter of probe tip, marking accuracy and examiner reliability all reportedly influence the outcome of probing^{1,2}. Measurements

may demonstrate significant variations between different clinicians and may also differ for the same clinician from one examination to another. It is important that high levels of reproducibility be achieved in clinical measurements.^{1,2} During the last decade various pressure-sensitive automated probes have been developed to improve reproducibility. Some authors have reported an improved reproducibility of probing measurements^{3,4}, whereas others have found no improvement in comparison with conventional probes.^{5,6,7,8} By the way, the most commonly used periodontal probes are still manual instruments with markings arranged in a variety of gradations. The main purpose of the present study was to evaluate intra- and inter-examiner reproducibility of probing depth measurements for a conventional manual probe (Williams).

Materials and Methods

Subjects

A total of 32 dental students in Tabriz Faculty of Dentistry volunteered to take part in the present study. The students, 17 females and 15 males, were 18-30 years old.

Each subject had at least 5 teeth, except the third molar, in the upper right quadrant. The teeth were well-aligned without evidence of severe malocclusion.

Probe

In this study, the periodontal probe used was a manual conventional probe (Williams style) with an 0.5-mm ball tip and markings at 1, 2, 3, 5, 8, 9 and 10 mm. All the measurements were rounded off to 0.5 mm.

Recordings

Each patient was examined by two examiners (two periodontists) in two sessions with an interval of 7-10 days. In each session, volunteers were examined by both examiners (examiners A and B). Therefore, 4 recordings were registered for each subject. To eliminate the possible bias, the order of examinations performed by 2 examiners was changed in a systematic

manner: 15 patients (A-B-B-A) and 17 patients (B-A-A-B).

Each tooth of the upper right quadrants, except the third molars, was examined. Probing depths were measured in 6 surfaces of each tooth (mesiofacial, midfacial, distofacial, mesiolingual, midlingual and distolingual). Probing depth measurements over 3 mm were excluded. A total of 218 teeth and 1295 surfaces were examined.

Data analysis

The paired-samples t-test was used to assess the mean differences between pairs of measurements for intra- and inter-examiner. P-values of <0.05 were defined as statistically significant.

Level of agreement between measurement pairs was calculated within defined limits of variations (in increments of ± 0.5 mm).

Results

The correlation coefficients between duplicate measurements for examiner A or B (intra-examiner) were $r = 0.730$ and $r = 0.814$, respectively. The corresponding value between measurements examined by A and B (inter-examiner) was $r = 0.627$. Although correlations between measurements were high, intra-examiner duplicate measurements showed higher correlation than inter-examiner ones.

Levels of agreement between duplicate measurements for intra-examiner A and intra-examiner B and also for inter-examiner A-B are presented in Table 1. Agreement of intra-examiner probing depth measurements at variations of ± 0.0 , ± 0.5 and ± 1.0 millimeters were 56.8, 89.0 and 99.6 percent, respectively. Agreement of inter-examiner measurements within these variations was 48.8, 84.2 and 98.5 percent.

Intra-examiner duplicate measurements for examiner A and examiner B showed no significant differences ($p=0.751$ and $p=0.783$, respectively), whereas differences between measurements made by examiner A and examiner B (inter-examiner differences) were statistically significant ($p=0.000$). Intra-examiner reproducibility was more than the inter-examiner one.

The same trend was consistently noted for two tooth types (anterior vs. posterior) and

also for 6 specified tooth surfaces. Table 2 compares the agreement between anterior and posterior teeth, while Table 3 shows a comparison between 6 specified tooth surfaces.

Agreement for duplicate probing depth measurements of anterior teeth was significantly higher than posterior teeth ($p < 0.05$). Variability in proximal surfaces (especially distal surfaces) was significantly more than mid-facial or mid-lingual surfaces ($p < 0.05$). The difference between the agreement of mesial and distal surfaces was not significant ($p < 0.422$). Facial surfaces showed significantly more agreement than lingual surfaces ($p < 0.05$).

Finally, measurement differences between the first and second examinations in each session were not significant ($p = 0.939$).

Discussion

The present study demonstrated that intra-examiner reproducibility was more than the inter-examiner one. This result confirms earlier findings.^{9,10} The possible reason may be the definite differences between several examiners, which include variations in the method of probing, force and so forth.^{11,12} However, agreement of duplicate measurements by means of the manual probe was found to be high within a range of 1 mm (above 98%). This is an acceptable tolerance limit for examiner agreement.⁸ This situation is important to consider. It suggests that in a clinical setting, a 1-mm level of disagreement between examiners, over time, may be acceptable for decision making. However, in the research setting, especially longitudinal studies where changes of 0.10 mm can demonstrate statistically significant values, a single examiner with more reliable instruments is recommended to minimize probing errors.^{7,8}

The correlation coefficients and agreements for intra- and inter-examiner duplicate measurements in this study were similar to most previous studies.^{6,13-16} However, the results were slightly higher than some of them. Two reasons could be mentioned: 1. highly-educated and well-trained examiners; 2. subjects with relatively

healthy periodontium (shallow probing depths result in better reproducibility⁶).

In the present study, anterior sites had better reproducibility than posterior sites. This finding coincides with earlier reports.^{5,6,15,17} Similarly, facial surfaces showed a slightly higher reproducibility than lingual surfaces and variability in proximal surfaces (especially distal surfaces) was more than mid-facial or mid-lingual surfaces. These findings are consistent with previous studies.^{6,17}

The greater reproducibility at anterior, facial and mid-facial/mid-lingual sites was not unexpected, on the grounds of accessibility and ease of probe alignment. Another reason is "searching phenomenon", described by Freed.¹² Regarding this issue, most examiners believe posterior and proximal sites have deeper probing depths. Therefore, they unconsciously probe these surfaces with more force. This may result in more variability and error.

There is concern that repeated probings of the same site would induce changes in the sulcus, influencing the measurements. One study showed that probing produces an instant but transient shock in gingival tissues. If a site is immediately re-probed, the second measurement could be influenced by the first. However, when a period of 5 minutes passes before the site is re-probed, the second measurement is not influenced by the first.¹⁸ None of the sites in our study were immediately re-probed. Consequently, no significant differences were observed between the first and second examinations in each session.

Finally, one advantage of the present study was the large study group. However, selecting subjects with normal periodontium may limit the use of these findings in periodontal patients with more probing depths.

Acknowledgements

The authors wish to extend their gratitude to all the staff in the Department of Periodontics in Tabriz Faculty of Dentistry.

Table 1. Agreement of duplicate probing depth measurements for intra-examiner A, intra-examiner B and inter-examiner A-B (%)

Variation (mm)	A	B	A-B
±0.0	55.6	57.9	48.8
±0.5	85.6	92.5	84.2
±1.0	99.3	99.9	98.5
±1.5	99.7	100.0	99.8
±2.0	100.0	100.0	100.0

Table 2. Agreement of duplicate probing depth measurements for intra-examiner A, intra-examiner B and inter-examiner A-B in anterior and posterior teeth (%)

Variation (mm)	Anterior		
	A	B	A-B
±0.0	59.3	60.9	47.8
± 0.5	89.7	94.9	86.2
± 1.0	100.0	100.0	99.4
± 1.5	100.0	100.0	100.0
±2.0	100.0	100.0	100.0
Variation (mm)	Posterior		
	A	B	A-B
±0.0	52.7	55.6	48.9
±1.0	98.8	99.9	97.9
±1.5	99.5	100.0	99.7
±2.0	100.0	100.0	100.0

Table 3. Agreement of duplicate probing depth measurements for intra-examiner A, intra-examiner B and inter-examiner A-B at 6 specified tooth surfaces (mesiofacial, midfacial, distofacial, mesiolingual, midlingual, distolingual) (%)

Variation (mm)	Mesiofacial		
	A	B	A-B
±0.0	52.1	46.3	43.2
±0.5	83.4	92.7	87.1
±1.0	99.5	100.0	99.1
±1.5	100.0	100.0	99.8
±2.0	100.0	100.0	100.0
Variation (mm)	Mid-facial		
	A	B	A-B
±0.0	69.6	75.1	65.9
±0.5	93.5	97.7	94.7
±1.0	97.7	100.0	98.6
±1.5	98.6	100.0	99.5
±2.0	100.0	100.0	100.0

Table 3. continued

Distofacial			
±0.0	45.4	52.3	40.2
±0.5	83.8	93.1	76.1
±1.0	99.5	100.0	97.9
±1.5	100.0	100.0	100.0
±2.0	100.0	100.0	100.0
Mesiolingual			
±0.0	46.1	55.0	41.2
±0.5	81.1	89.4	80.4
±1.0	100.0	100.0	98.4
±1.5	100.0	100.0	100.0
±2.0	100.0	100.0	100.0
Mid-lingual			
±0.0	65.3	69.6	58.5
±0.5	93.0	98.2	90.0
±1.0	99.5	100.0	99.3
±1.5	100.0	100.0	100.0
±2.0	100.0	100.0	100.0
Distolingual			
±0.0	55.2	49.1	41.4
±0.5	78.8	84.1	76.7
±1.0	99.5	99.5	97.9
±1.5	100.0	100.0	99.5
±2.0	100.0	100.0	100.0

Table 4. Agreement of duplicate probing depth measurements for intra-examiner A, intra-examiner B and inter-examiner A-B for facial and lingual surfaces (%)

Variation (mm)	Facial		
	A	B	A-B
±0.0	55.7	57.9	49.8
±0.5	86.9	94.5	85.9
±1.0	98.8	100.0	98.5
±1.5	99.5	100.0	99.8
±2.0	100.0	100.0	100.0
Lingual			
±0.0	55.5	57.9	47.0
±0.5	84.3	90.6	82.4
±1.0	99.7	99.8	98.5
±1.5	100.0	100.0	99.8
±2.0	100.0	100.0	100.0

Table 5. Agreement of duplicate probing depth measurements for intra-examiner A, intra-examiner B and inter-examiner A-B for mesial, mid and distal surfaces (%)

Variation (mm)	Mesial		
	A	B	A-B
±0.0	49.1	50.7	42.2
±0.5	82.3	91.0	83.8
±1.0	99.8	100.0	98.8
±1.5	100.0	100.0	99.9
±2.0	100.0	100.0	100.0
	Mid		
±0.0	67.5	72.6	62.2
±0.5	93.3	97.9	92.6
±1.0	98.6	100.0	99.0
±1.5	99.3	100.0	99.8
±2.0	100.0	100.0	100.0
	Distal		
±0.0	50.3	50.7	40.8
±0.5	81.3	88.6	76.3
±1.0	99.5	99.8	97.9
±1.5	100.0	100.0	99.7
±2.0	100.0	100.0	100.0

References

- Newman MG, Takei HH, Carranza FA. *Carranza's Clinical Periodontology*, 9th ed. Philadelphia: Saunders; 2002: 441-53, 489-91.
- Lindhe J. *Clinical Periodontology and Implant Dentistry*, 4th ed. Copenhagen: Blackwell; 2003: 407-9.
- Rams TE, Slots J. Comparison of two pressure-sensitive periodontal probes and a manual periodontal probe in shallow and deep pockets. *Int J Periodontics Restorative Dent* 1993; 13: 520-9.
- Magnusson I, Fuller WW, Heins PJ, Rau CF, Gibbs CH, Marks RG, Clark WB. Correlation between electronic and visual readings of pocket depths with a newly developed constant force probe. *J Clin Periodontol* 1988; 15:180-4.
- Mayfield L, Bratthall G, Attstrom R. Periodontal probe precision using 4 different periodontal probes. *J Periodontol* 1996; 23: 76-82.
- Wang SF, Leknes KN, Zimmerman GJ, Sigurdsson TJ, Wikesjo UME, Selvig KA. Reproducibility of periodontal probing using a conventional manual and an automated force controlled electronic probe. *J Periodontol* 1995; 66: 38-46.
- Khocht A, Chang KM. Clinical evaluation of electronic and manual constant force probes. *J Periodontol* 1998; 69: 19-25.
- Perry DA, Taggart EY, Leung A, Newbrun E. Comparison of a conventional probe with electronic and manual pressure-regulated probes. *J Periodontol* 1994; 65: 908-13.
- Hassell TM, German MA, Saxer UP. Periodontal probing: Inter-investigator discrepancies and correlations between probing force and recorded depth. *Helv Odontol Acta* 1973; 17: 38.
- Buduneli E, Aksoy O, Kose T, Atilla G. Accuracy and reproducibility of two manual periodontal probes. *J Clin Periodontol* 2004; 31: 815.
- Listgarten MA. Periodontal probing: What does it mean? *J Clin Periodontol* 1980; 7: 165-76.
- Freed HK, Gapper RL, Kalwarf KL. Evaluation of periodontal probing forces. *J Periodontol* 1983; 54: 488-92.
- Osborn J, Stoltenberg J, Huso B, Aeppli D, Pihlstrom B. Comparison of measurement variability using a standard and constant force periodontal probe. *J Clin Periodontol* 1990; 61: 497-503.

14. Kingman A, Loe H, Anerud A, Boysen H. Errors in measuring parameters associated with periodontal health and disease. *J Periodontol* 1991; 62: 477-86.
15. Mullally BH, Linden GJ. Comparative reproducibility of proximal probing depth using electronic pressure-controlled and hand probing. *J Clin Periodontol* 1994; 21: 284-8.
16. Wang SF, Leknes KN, Zimmerman GJ, Sigurdsson TJ, Wikesjo VME, Selvig KA. Intra- and inter-examiner reproducibility in constant force probing. *J Clin Periodontol* 1995; 22: 918-22.
17. Badersten A, Nilveus R, Egelberg J. Reperducibility of attachment level measurements. *J Cin Priodontol* 1984; 11: 475-85.
18. Clark B, Yang M, Magnusson I. Measuring clinical attachment: reproducibility of relative measurements with an electronic probe. *J Periodontol* 1992; 63: 831-8.