Probing around teeth

Barendregt, D.S.

Publication date
2009

Citation for published version (APA):
COMPARISON OF 2 AUTOMATED PERIODONTAL PROBES AND 2 PROBES WITH A CONVENTIONAL READOUT IN PERIODONTAL MAINTENANCE PATIENTS

CHAPTER 7

D. S. Barendregt
U. van der Velden
M.F. Timmerman
G.A. van der Weijden

Department of Periodontology, Academic Centre for Dentistry Amsterdam, The Netherlands
A periodontal probe till today remains the most important diagnostic tool in periodontal diseases. It is used to establish the presence and severity of the disease and also used to assess the effect of periodontal treatment. The probe enables the clinician to determine pocket depth, attachment level, presence of plaque and calculus, and anatomical features of the root. Reliable measurements of the pocket depth and the attachment level are critical to both longitudinal clinical studies and routine clinical assessment in periodontal therapy. Current probing methods are subject to various errors e.g. measurement outcomes are strongly dependent on probing force (Hassell et al. 1973; Mombelli et al. 1992; Van der Velden 1979). Therefore variations in probing force are appear to be evident between different examiners but also for a single examiner (Gabathuler et al. 1971). When measuring a pocket the degree of probe tip penetration is also influenced by the presence of inflammation. Even with relatively high forces the probe tip usually fails to reach the connective tissue attachment in healthy sites (Fowler et al. 1982). With low probing pressures, the probe tip generally stops at the level of intact connective tissue fibers or beyond in deep inflamed sites (Bulthuis et al. 1998).

During the last decades various pressure-sensitive automated probes have been developed to reduce the factor of variability of probing force (Chamberlain et al. 1985; Garnick et al. 1989). Some authors have reported an improved reproducibility of probing measurements (Abbas et al. 1982; Osborn et al. 1990; Walsh et al. 1989; Wang et al. 1995), whereas others found no improvement of the reproducibility when using constant force probes (Kalkwarf et al. 1986; Quirynen et al. 1993; Van der Velden et al. 1980; Watts 1987).

The Florida Probe® introduced by Gibbs et al. (1988) has shown to be more reproducible than manual probing in a number of studies (Gibbs et al. 1988; Magnusson et al. 1988; Yang et al. 1992). At present this probe is considered the “golden standard” of the automated probes based on the extensive research on the validity of the Florida Probe® (Grossi et al. 1996; Osborn et al. 1992; Reddy et al. 1997). Also the Brodontic® probe, with a spring loaded hinge handle, was developed (Borsboom et al. 1981) to overcome the problem of varying probing forces. This probe showed a better reproducibility of probing depth measurements than a manual probe (Simons et al. 1987). Through its simple design this probe is an attractive solution to control probing pressures in daily practice and in field studies. Several studies have used this probe to ensure a constant pressure (Barendregt et al. 1996; Breen et al. 1997; Timmerman et al. 2000) However, up to now, it has not been compared to other pressure probes such as the Florida Probe®. More recently a new probe was developed in the
Netherlands (Jonker Probe®) (Fig 2). The design of this automated probe is based on the constant force probe developed by (Van der Velden 1978). Like the Florida Probe® it has an electronic readout. The electronic recording of the measurements offers advantages for operators who work alone and it eliminates scribe errors in clinical research.

Periodontal probing in patients with untreated periodontal disease might also be influenced by remaining calculus, plaque and over contouring of restorations. In order to minimize this problem Wang et al. (1995) selected patients in the maintenance phase. When testing for reproducibility, these subjects with relatively healthy reduced periodontium provide sites in which optimal probe angulation (Karim et al. 1990; Watts 1989) and positioning (Karim, et al. 1990) can be facilitated.

The aim of the present study was to test in periodontal maintenance patients whether the systems for pressure control that have been commercially developed contribute to more reproducible probing depth measurements as compared to a manual probe.

Material and Methods

Patients

In total 12 periodontal maintenance patients were selected for the study. They had an initial diagnosis of moderate to advanced periodontitis, on the basis of manual probing depth measurements and radiographs. All sites had received initial periodontal therapy consisting of instruction in plaque control measures, supra-/subgingival debridement and periodontal surgery when needed. Following the active treatment they were enrolled in a 3 to 4 monthly maintenance protocol. In each patient 4 teeth (preferably first molars) showing at least at one site a pocket of ≥ 5mm, were included in the study. For each selected tooth 6 sites were recorded which resulted in 288 sites available for the study. These experimental teeth were equally distributed between the arches and incorporated shallow (<4mm), moderate (≥4 and <7mm) and deep sites (≥7mm). In addition at screening and selection, the level of gingival inflammation was evaluated through recording of the presence or absence of bleeding on manual probing (BOP).
**Description of Probes used**

a) *Florida Probe*® (Florida Probe Company, Gainesville Florida, USA) was equipped with a tapered tine with a diameter of 0.4 mm at the tip increasing to 0.5 mm at the 5 mm and 0.6 mm at the 10 mm marking (Fig. 1). The probing force was adjusted according to the manufacturer’s guidelines to 0.20 N resulting in a probing pressure of 159 N/cm² (Gibbs et al. 1998).

![Fig. 1](image1)

b) *Jonker Probe*® (*Jonkers Data, Staphorst, Netherlands*); It has a tapered tine with a diameter at the tip of 0.5 mm increasing to 0.6 mm at the 5 mm and 0.7 mm at the 10 mm marking (Fig. 2). The probing force of Jonker Probe was 0.30 N, achieving a probing pressure of 153 N/cm².

![Fig. 2](image2)

c) *Brodontic probe* (Prima, Byfleet, England, UK) with Williams markings; this probe has a spring loaded hinged handle exerting a constant force. A tapered tine (Ash Dentsply, Weybridge, Surrey, England, UK) was mounted with a diameter at the tip of 0.5 mm increasing to 0.6 mm at the 5 mm and 0.7 mm at the 10 mm marking (Fig. 3). Based on Van
der Velden et al. (1979), who used 240 N/cm² in maintenance patients, a probing force of 0.50 N was used to achieve a comparable probing pressure (255 N/cm²).

**Fig. 3**

d) Conventional manual probe (Hu-Friedy, Chicago, USA) with Williams markings (Fig.4); this probe had a tapered tine with a diameter of 0.5 mm at the tip increasing to 0.6 mm at the 5 mm and 0.7 mm at the 10 mm marking.

**Fig. 4**

**Probing depth measurements**

Duplicate recordings were made with a 30 minutes interval at the distobuccal (DB), midbuccal (B), mesiobuccal (MB), distolingual (DL), lingual (L) and mesiolingual (ML) sites at the 4 experimental teeth in each patient. This amounted to 288 evaluable sites. The recordings were performed at the first visit (Day 0) and again at the second visit 1 week later (1 Week). In each patient, each of the 4 selected experimental teeth was assigned to a random probing order according to a Latin square design. In all 6 sites of each experimental tooth, all 4 probes were used based in this assigned random order. This
order remained the same for each individual tooth for the duplicate recording and the recordings one week apart. In order to minimize the effect of bias as a result of intra-examiner reproducibility, 2 examiners were chosen (MP and YIJ). Each examiner was assigned 6 patients and consequently 144 sites.

The probes were inserted parallel to the root surface and directed apically toward the perceived location of the apex of the root. When the pre-set force was reached (in case of the 3 pressure probes) the probing depth was recorded and stored by the computer software or when appropriate written down on a case record form by the assistant. With the manual probe and the Brodontic® probe the recordings were rounded off to the nearest whole millimeter.

Data analysis

Analysis of probing measurements for the different probes was performed using the site as the unit of measurement repeats. Differences between probes, duplicates and week were tested by use of a mixed model analysis of variance corrected for examiner and patient effects. For differences between repeats and between Day 0-1 week the Standard Error (SE) and the 95% Confidence Interval (CFI) were calculated. Furthermore correlation coefficients were calculated for all probes over all sites comparing Day 0 and 1 Week measurements. To test for systematic differences between sessions paired Student t-test were used. p-values of <0.05 were accepted as statistically significant.

Results

Overall 96% first molars and 4% second premolars were evaluated. Table 1 shows the mean probing depths at screening and selection with the manual probe at site level. The mean probing depth over the 288 sites was 3.90 mm. The level of gingival inflammation, as assessed by bleeding on probing, was 21.5%. The mean results were subdivided into shallow (0-3.5 mm) moderately deep (4-6.5 mm) and deep sites (7-10 mm). The proportion of shallow group was 51%, moderate deep sites 41% and deep sites 8%. Compared to the buccal/lingual sites the mean probing depth at the approximal surfaces was higher up to the level of 4.54 mm and the proportion of moderately deep pockets was also higher (59%). The buccal/lingual sites offered primarily shallow pockets (90%) and with a mean probing
depth of 2.62 mm. The level of bleeding on probing at the approximal sites was 23% and 16% at the buccal/lingual surfaces.

Table 1. Mean probing depths (mm), bleeding on probing (%) and shallow(S)-moderate(M)-deep(D) sites (%) based on screening measurements (Williams probe) over all sites, approximal and buccal/lingual.

<table>
<thead>
<tr>
<th></th>
<th>Mean (SD)</th>
<th>BOP</th>
<th>S</th>
<th>M</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients (n=12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All sites</td>
<td>3.90 (1.65)</td>
<td>21.5</td>
<td>51%</td>
<td>41%</td>
<td>8%</td>
</tr>
<tr>
<td>Approximal</td>
<td>4.54 (1.40)</td>
<td>23</td>
<td>30%</td>
<td>59%</td>
<td>11%</td>
</tr>
<tr>
<td>Buccal / lingual</td>
<td>2.62 (0.90)</td>
<td>16</td>
<td>90%</td>
<td>10%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Table 2 shows the mean overall results with the 4 probes. The mean probing depth as established with the Florida Probe® (FP) and the Jonker® Probe (JP) did not differ (3.33 vs. 3.32), neither was there a difference between the Brodontic® probe (BP) and the manual probe (MP) (3.95 vs. 3.93 mm). The FP and JP measured a significantly lower mean probing depth than the BP and the MP. When subdividing the measurements into approximal and buccal/lingual sites, similar observations were made.

Table 2. Overall mean probing depths for all probes over all sites, approximal and buccal/lingual sites.

<table>
<thead>
<tr>
<th></th>
<th>All sites n = 288</th>
<th>Approx n = 192</th>
<th>Buccal/lingual n = 96</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Florida Probe®</td>
<td>3.33&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>1.42</td>
<td>3.91&lt;sup&gt;a,b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Jonker Probe</td>
<td>3.32&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>1.36</td>
<td>3.82&lt;sup&gt;a,b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Brodontic probe</td>
<td>3.95</td>
<td>1.52</td>
<td>4.59</td>
</tr>
<tr>
<td>Manual Probe</td>
<td>3.93</td>
<td>1.59</td>
<td>4.61</td>
</tr>
</tbody>
</table>

<sup>a</sup> significant difference with the Brodontic probe p< 0.05
<sup>b</sup> significant difference with the Manual probe p< 0.05
Duplicate measurements (session 1 & 2) at Day 0 and at 1 Week are shown in Table 3. The FP, the BP and the MP show no significant differences between the duplicate measurements. With the JP however at Day 0 and 1 Week, the second measurement was deeper and increased with 0.19 and 0.14 mm respectively ($p=0.01$).

**Table 3. Mean values and differences of duplicate measurements at Day 0 and 1 Week (session 1 and 2)**

<table>
<thead>
<tr>
<th>Probe Type</th>
<th>Day 0 session1 Mean</th>
<th>Day 0 session2 Mean</th>
<th>Diff 1-2</th>
<th>SE</th>
<th>95% CFI lower</th>
<th>95% CFI upper</th>
<th>1 Week session1 Mean</th>
<th>1 Week session2 Mean</th>
<th>Diff 1-2</th>
<th>SE</th>
<th>95% CFI lower</th>
<th>95% CFI upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Florida Probe®</td>
<td>3.29</td>
<td>3.41</td>
<td>-0.12</td>
<td>0.06</td>
<td>-0.24</td>
<td>0.00</td>
<td>3.32</td>
<td>3.29</td>
<td>0.03</td>
<td>0.06</td>
<td>-0.09</td>
<td>0.16</td>
</tr>
<tr>
<td>Jonker Probe®</td>
<td>3.22</td>
<td>3.41*</td>
<td>-0.19</td>
<td>0.06</td>
<td>-0.30</td>
<td>-0.07</td>
<td>3.25</td>
<td>3.38*</td>
<td>-0.14</td>
<td>0.05</td>
<td>-0.24</td>
<td>-0.03</td>
</tr>
<tr>
<td>Brodonti Probe®</td>
<td>3.90</td>
<td>3.96</td>
<td>-0.06</td>
<td>0.04</td>
<td>-0.13</td>
<td>0.10</td>
<td>3.93</td>
<td>4.02</td>
<td>-0.08</td>
<td>0.03</td>
<td>-0.13</td>
<td>0.01</td>
</tr>
<tr>
<td>Manual Probe®</td>
<td>3.91</td>
<td>3.93</td>
<td>-0.02</td>
<td>0.04</td>
<td>-0.09</td>
<td>0.05</td>
<td>3.92</td>
<td>3.95</td>
<td>-0.03</td>
<td>0.03</td>
<td>-0.09</td>
<td>0.04</td>
</tr>
</tbody>
</table>

When comparing mean probing depths at Day 0 and 1 Week (first session only) no significant differences for the duplicate assessment 1 week apart for any of the 4 probes were found (Table 4). The correlation coefficients between the first assessment at Day 0 and 1 Week are presented in the last column of Table 4. The automated probes show a comparable value, for FP and for the JP (0.76 and 0.75 resp.). The BP and the MP show higher correlation coefficients between sessions of 0.90 and 0.89 respectively.
Table 4. Differences and correlation coefficient at Day 0-1 Week (session 1) for all sites.

<table>
<thead>
<tr>
<th></th>
<th>Difference</th>
<th>SE</th>
<th>95% CFI</th>
<th>Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>lower</td>
<td>upper</td>
</tr>
<tr>
<td><em>Florida Probe</em></td>
<td>-0.03</td>
<td>0.06</td>
<td>-0.15</td>
<td>0.09</td>
</tr>
<tr>
<td><em>Jonker Probe</em></td>
<td>-0.03</td>
<td>0.06</td>
<td>-0.15</td>
<td>0.10</td>
</tr>
<tr>
<td><em>Brodontic probe</em></td>
<td>-0.04</td>
<td>0.04</td>
<td>-0.15</td>
<td>0.01</td>
</tr>
<tr>
<td><em>Manual Probe</em></td>
<td>-0.02</td>
<td>0.05</td>
<td>-0.11</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Fig. 5 a, b, c, d illustrates the frequency distribution of the differences between the measurements at Day 0 and 1 Week (1<sup>st</sup> session) stratified in the categories shallow (<4 mm), moderate (≥4 and <7 mm) and deep sites (≥ 7 mm). The range of differences between 2 assessments for the probes over all sites with a conventional readout (BP and MP) does not exceed between -2 and 2 mm. The range for the FP is -3 to 3 mm and for the JP even -5 to 5 mm. In 48-59% of the assessments no differences was observed between for the 4 probes. The measurement error between -1 and 1 mm is found in 94% of the cases with the BP and 96% with the MP. In comparison the FP showed in 84% of the measurements between -1 and 1 mm and the JP 88%.
Figure 5 a

Differences all sites

Figure 5 b

Differences shallow sites (n = 144)
Discussion

In this study 3 commercially available pressure probes and a manual probe were tested for their ability to provide reproducible measurements. A good range in probing depths should be available in order to compare these different probes in a proper way. As a result of the inclusion criteria used in this study, pockets of \( \geq 4 \) mm amounted to 49%. In comparison Wang et al. (1995) presented their data in maintenance patients on 20% of pockets \( \geq 4 \) mm. Mayfield et al. (1996) based their conclusions on a percentage of approximately 17% of pockets of \( \geq 5 \) mm whilst in the present study the percentage amounts to 36%. Reddy et al. (1997) presented data including 54% pockets of \( \geq 5 \) mm. However their sites were obtained in untreated moderate to advanced periodontitis patients. Taken into account the aforementioned studies, the present study appears to have a clinically representative study population providing a wide range of probing depths (Table 1). It is therefore feasible to assume that the results in this study are applicable for measurements of probing depths in periodontal maintenance patients.

The influence of probing pressures on the results of probing depth measurements has been studied by several authors (Hassell et al. 1973, Van der Velden 1979, Mombelli et al. 1992). The FP and the JP used pressures of 159 N/cm\(^2\) and 153 N/cm\(^2\) respectively whereas the BP was set to a pressure of 255 N/cm\(^2\). Since no differences in mean probing depths were found between the MP and the BP (Table 2) it is likely that a comparable probing pressure was exerted by the MP. The difference in probing pressures could therefore explain the higher mean probing depths with the BP and MP as compared to the FP and JP. Since it has been shown that the periodontal condition greatly influences the results of probing depth measurements (Bulthuis, et al. 1998; Fowler, et al. 1982; Garnick, et al. 1989; Hancock et al. 1981; Polson 1980; Robinson et al. 1979; Van der Velden 1980), in the present study all comparisons for the 4 probes were made at the same sites thereby controlling for the influence variations in periodontal inflammation. Based on the studies where different probing pressures were used to evaluate the true attachment level with the aforementioned probing systems (Van der Velden 1979; Fowler et al. 1982; Bulthuis et al. 1998) it is likely that the probe tip of the 2 automated probes employing relatively low probing pressures, did not reach the bottom of the pocket in the present study population of relatively healthy maintenance patients. Further research is needed to establish the possible consequences for the prediction of the long term stability of the periodontal condition on the basis of probing depth measurements with relatively low or high probing pressures.
The results of the present study show that the reproducibility of the automated force controlled probes is somewhat lower compared to the BP and MP. When comparing the probing measurements from Day 0 and 1 week differences of 2 mm and more were found for the automated probes and especially the JP (Figure 5 a,b,c,d). Bulthuis et al. (1998) showed that in inflamed situations low probing pressures are sufficient in order to reach the histological bottom of the pocket. However these authors also suggested that, when using light forces the examiner may run the risk of not entering the orifice of the pocket with the probe tip. This would especially be the case in more healthy sites where the gingival margin lies tight around the neck of the tooth. Interestingly in the present study Figure 5b shows a difference of 3 mm for the JP and the FP in the shallow group where the range of pocket depths lies between 0 and 3.5 mm (mean pocket depth of 2.62 mm). This suggests that at some sites these particular probes did not enter the pocket. This phenomenon could easily explain why the BP and the MP show less variability since they use higher pressures. This suggestion is consistent with the findings of Wang et al. (1995). In their study with periodontal maintenance patients, the manual probe also proved to be more reproducible than the automated force controlled probe. The pressure of this probe was set to 156 N/cm² comparable to the pressure used for the FP and JP in the present study. Waerhaug (1952) suggested that in more healthy situations, in order to reach the “true” bottom of the pocket with light forces, a thinner probe is needed. Since the FP has a smaller diameter at the tip of the probe than the JP, it may be more capable of entering the pocket with a healthy marginal gingiva. This could explain the larger variation of measurements between Day 0 and 1 Week with the JP as compared to the FP. Another explanation for the lower reproducibility of the automated probes could be the bulky anterior part of the electronic probes. This does make it difficult to get adequate access to posterior probing (Wang et al. 1995). The reproducibility of the JP may also be influenced by the total length of the tine of the probe and the sleeve in which it runs (Fig. 2). With increasing dimensions of this measuring device it will become more difficult to reach the posterior sites of the dentition. Comparing lengths of the MP (16 mm) and the FP (22 mm) with the JP (32 mm) it is evident that the greater dimensions of the JP will make it more difficult to manipulate inside the mouth of a patient. This suggestion is supported by the finding that the differences between repeated measurements with a magnitude of 3 mm and more were found mainly on the disto-lingual molar sites with both the JP and FP. The greater dimensions of JP may have also been responsible for the finding that on Day 0 and 1 Week the JP obtained deeper probing depth measurements at the second probing procedure. At the second probing the examiner may have remembered the difficulties
in putting the probe in the right position. Such an increase in deeper measurements at the second probing procedure was not found for the FP, BP and MP. This lack in differences between duplicate measurements seems to be in contrast to the literature since Janssen et al. (1988) observed deeper measurements in a second, third and fourth session on the same day using a force-controlled probe (approximately 240 N/cm²). Therefore Janssen et al. (1988) suggested that 2 duplicate measurements should be averaged to get a more accurate score. Based on the data presented in this study this does not appear to be necessary since the expected deeper second measurement (Janssen et al. 1988) was not observed with the FP, the BP and the manual probe in this study.

It has been generally accepted in the literature that a difference between two probing measurements with a time interval of 2 months or more, that exceeds 3 times the standard deviation (0.80-0.92 mm in the study of Badersten et al. 1984) is the result of disease or therapy. For all 4 probes a high proportion of the measurements showed a difference within -1 and 1 mm (84 to 96%). When these data are compared to other studies like Badersten et al. (1984), Watts et al. (1987), Mayfield et al. (1996) and Breen et al. (1997), the present data are within the same range. The Brodontic® probe and the manual probe showed a high correlation coefficient (0.83-0.93) between repeated examinations (Table 4). This corresponds with the data presented by Mayfield et al. (1996) where the manual probe, in a comparable protocol and patient group, showed similar correlation coefficients. Some studies have reported that constant force probes are more accurate (Gibbs, et al. 1988; Walsh, et al. 1989) whilst other studies prefer the manual probe (Mayfield et al. 1996; Osborn, et al. 1990; Wang, et al. 1995)). The present study has shown that the manual probe in the hands of the experienced examiner has a good reproducibility.

In conclusion, the automated probes showed a lower level of reproducibility than the BP and MP. In addition, due to the lower probing pressure of the FP and JP these probes measured less deep as compared to the BP and MP. It is suggested that a pressure level of approximately 250 N/cm² may be needed in periodontal maintenance patients since otherwise in a number of cases the probe fails to enter the pocket. Since in this study a good reproducibility was achieved and similar probing depth recordings were obtained as compared to the pressure controlled BP, the manual probe still remains a reliable tool in daily periodontal practice.
Acknowledgements

The authors wish to thank Y. IJzerman and M. Piscaer for their effort and perseverance in facilitating this study in patient selection and data collection.

References


