Published in final edited form as:

J Clin Periodontol. 2012 November ; 39(11): 1032–1041. doi:10.1111/j.1600-051X.2012.01941.x.

Effects of different manual periodontal probes on periodontal measurements

Birte Holtfreter¹, Dietrich Alte², Christian Schwahn³, Moïse Desvarieux^{4,5,6}, and Thomas Kocher¹

¹Unit of Periodontology, Department of Restorative Dentistry, Periodontology, and Endodontology, University Medicine, Ernst Moritz Arndt University of Greifswald, Greifswald, Germany

²Department of SHIP/Clinical-Epidemiological Research, Institute for Community Medicine, University Medicine, Ernst Moritz Arndt University of Greifswald, Greifswald, Germany

³Department of Prosthodontics, Gerodontology, and Biomaterials, University Medicine, Ernst Moritz Arndt University of Greifswald, Greifswald, Germany

⁴Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, USA

⁵Institut National de la Santé et de la Recherche Médicale, Unité 738, Paris, France

⁶École des Hautes Études en Santé Publique, Paris et Rennes, France

Abstract

Aim—To quantify the digit preference effect for three manual periodontal probes and to calculate correction values to enable comparison of studies with equal recording protocols, but different periodontal probes.

Material and Methods—A prospective in vivo crossover study was conducted with a six-sequence three-period design. Six examiners assessed attachment loss (AL), probing pocket depth (PD) and gingiva height (GH) at four surfaces, full-mouth, in six generally healthy subjects using three manual probes: PCP11 (3-3-3-2 mm increments), PCP2 (2 mm increments), and PCPUNC15 (1 mm increments).

Results—Distributions of AL, PD and GH differed between probes ($p < 0.001$). Compared with PCPUNC15, periodontal measurements coinciding with probe markings of PCP11 and PCP2, respectively, were preferentially named by examiners. Digit preference was most pronounced for PD, but less for AL and GH. In multilevel models, PD differed significantly between all three

© 2012 John Wiley & Sons A/S

Address: Birte Holtfreter, Unit of Periodontology, Department of Restorative Dentistry, Periodontology, and Endodontology, University Medicine, Ernst Moritz Arndt University of Greifswald Rotgerberstr. 8, 17475 Greifswald, Germany, birte.holtfreter@uni-greifswald.de.

Conflict of interests

There are no conflicts of interests associated with this work.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1: Multilevel model evaluating the effect of probe, period, examiner and pairwise interactions on three periodontal measurements.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

probes ($p < 0.05$); probe- and examiner-related effects were also observed for AL and GH. Correction values for pairwise combinations of probes were determined.

Conclusions—We provided empirical evidence and quantified the effect of probe type on periodontal measurements. Differences in probe type should be considered when comparing periodontal data within and between epidemiological studies and appropriate corrections, provided here, should be applied.

Keywords

bias correction values; digit preference effect; manual periodontal probe; periodontal attachment loss; probing pocket depth

In epidemiological studies, sequential probing of periodontal pockets is the most commonly used method to quantify the periodontal status based on attachment loss (AL) and probing pocket depth (PD) values and to measure change or differences in the periodontal status in clinical, longitudinal or between cross-sectional studies (Listgarten 1980, Hefti 1997). However, variation in methodologies including periodontal recording protocols (Susin et al. 2005, Beck et al. 2006, Vettore et al. 2007) limits the opportunity for comparisons of epidemiological studies. Additional factors affecting inter- and intra-examiner variability and measurement accuracy for periodontal measurements include the experience and number of examiners (Grossi et al. 1996, Reddy et al. 1997), lighting conditions, the degree of periodontal destruction in examined subjects (Fleiss et al. 1991), severity of tissue inflammation (Listgarten 1980, Hefti 1997, Garnick & Silverstein 2000), type of teeth (Grossi et al. 1996), sites of measurement (Grossi et al. 1996, Reddy et al. 1997) and the choice of periodontal instruments regarding shape and tip diameter (Atassi et al. 1992, Garnick & Silverstein 2000). Errors inherent to the use of a periodontal probe include variation in probing force (Reddy et al. 1997, Garnick & Silverstein 2000, Larsen et al. 2009), probe placement and angulation (Persson 1991), visual or tactile difficulty in detecting the cemento- enamel junction (CEJ) (Watts 1987, Reddy et al. 1997) and erroneous recording of measurements (Hefti 1997, Garnick & Silverstein 2000, Silva-Boghossian et al. 2008).

The digit preference effect (Thavarajah et al. 2003, Camarda et al. 2008) might be a further aspect influencing reliability and accuracy of probing measurements, having to do with rounding error associated with specific instruments. Digit preference was reported in measurements in other fields, such as self-reported year of menopause (Crawford et al. 2002), self-reported height and weight (Rowland 1990), or blood pressure measurements (Camarda et al. 2008). For example, for manual diastolic blood pressure measurements, values complying with the rough graduation of the blood pressure metre, that is, 80, 85 or 90, are named more often than values of the more precise graduation, that is, 81, 82 or 83. Digit preference can bias the observed distribution of values (Camarda et al. 2008).

Applying this concept to periodontal probing, periodontal landmarks (gingiva margin and/or CEJ) and probe markings aid decisions on periodontal assessments. Systematic rounding errors might occur in dividing regions between horizontal lines (Kashmere & Kirk 1997). AL measurements differ from PD or gingiva height (GH) measurements in at least one aspect: with an exposed and visible CEJ, the two probe graduation markings below and above the CEJ are available to estimate AL. To measure PD or GH, only one probe graduation marking above the gingival margin is visible and supports decision of the examiner. Until now, the effect of probe graduation through digit preference has not been described in the periodontal literature. As the choice of the graduation markings instrument might have a substantial influence on periodontal measures in clinical and epidemiological

studies, there is a need for the quantification of the digit preference effect for manual probes with different graduations.

As probes with different graduation markings were used in the baseline and follow-up waves of the Study of Health in Pomerania (Hensel et al. 2003, Volzke et al. 2011), bias correction values accounting for digit preference effects are required. Recently, Larsen et al. (2009) provided correction factors that compensate for the different probing pressures used. Analogous, the use of correction values might be reasonable to enable comparison of periodontal measurements between studies with different periodontal probes but identical recording protocols.

The aim of our study was to (i) quantify the digit preference effect for three manual probes utilized in the Study of Health in Pomerania (SHIP) and supplied by one of the major manufacturers (Hu-Friedy, Chicago, IL, USA) and (ii) to provide bias correction values for three clinical assessments for pairwise combinations of those probes.

Material and Methods

Study design and subjects

This prospective study was conducted as an in vivo crossover study with a six-sequence three-period design. Six examiners recorded AL, PD and GH in six generally healthy subjects with three different manual probes (full factorial design). All measurements were conducted on three non-consecutive days within 2 weeks with a wash-out period between examination days of at least 2 days. On each day, every examiner recorded periodontal measurements in each of the six subjects in random order. Altogether, each examiner performed six examinations per day. For each examiner, the order of probes and subjects was randomly assigned. On each day, each subject was examined once by each of the six examiners, while always two examiners used the same probe. Thus, each subject was examined six times each day, resulting in 18 examinations in total per subject (six examiners $\times$ three probes). Overall, 54 periodontal examinations were conducted. The study protocol was approved by the local ethics committee/institutional review board.

Subjects selected were volunteers from a pool of subjects used for calibration studies and pre-tests for SHIP. They were generally healthy with an age range of 42–55 years; five subjects were male.

Periodontal measurements

Attachment loss, PD and GH values were recorded using a full-mouth design (excluding third molars) with four sites per tooth (distobuccal, mesiobuccal, midbuccal and midlingual/midpalatal). PD equals the distance from free gingival margin (FGM) to pocket base. If the FGM was located apical to the CEJ, AL was measured as the distance between CEJ and pocket base; no GH was recorded. If the FGM was located coronal to the CEJ, only the distance from CEJ to FGM (gingiva height; GH) and PD was recorded; AL was calculated as PD – GH. Where the determination of the CEJ was indistinct (wedge-shaped defects, fillings and crown margins), only PD was recorded. Measurements were recorded in whole millimeters.

All examiners were experienced dentists, with five of them being calibrated examiners in SHIP (Hensel et al. 2003). Periodontal measurements were assessed using three different manual periodontal probes, which differ with respect to scale, diameter, shape and graduation quality: PCP11, PCP2 and PCPUNC 15 (Hu-Friedy). The PCP11 probe had markings at 3, 6, 8 and 11 mm (Fig. 1a,d). For PCP2, dark bands divided the probe at 2, 4, 6, 8, 10 and 12 mm (Fig. 1b,e). The PCPUNC15 had coloured and engraved markings at 1 mm

intervals (Fig. 1c,f). Tip diameter for PCP11, PCP2 and PCPUNC15 was 0.48, 0.50 and 0.57 mm respectively. Probing force was trained using a scale; target probing force was 0.5 N.

Statistical analyses

Statistical analyses were performed with STATA/SE 10.0 (StataCorp 2007) and R 2.13.0 (R Development Core Team 2011). Data are presented as mean with 95% confidence interval or numbers (percentages). A p -value < 0.05 was considered statistically significant. The distribution of periodontal measurements between periodontal probes was compared with the chi-squared test, ignoring the hierarchical structured of the data.

To detect digit preference patterns of PCP2 and PCP11 compared with PCPUNC15, which was considered the gold standard periodontal probe because of its fine and equidistant scaling, distance dimensions for two superimposed frequency distributions were considered. Corresponding graphs plot differences in relative frequencies between PCP2/PCP11 and PCPUNC15 (Fig. 2). If a certain measurement value, for example, PD of 3 mm, was named more often for PCP2/PCP11 than for PCPUNC15, differences were negative. If a certain measurement value was named more often for PCPUNC15 than for PCP2/PCP11, differences were positive.

Data for periodontal measurements were analyzed using multilevel regression models. Random effects accounted for the hierarchical structure; that is, subjects within sequence, teeth within subjects and sites within teeth. Fixed effects included period and periodontal probe. Terms for examiner, the interaction between probe and period, the interaction between probe and examiner, and the carry-over effect were included in the final model if $p_{\text{overall}} < 0.10$. Assumptions including normally distributed, inconspicuous residuals (residual plots) were met.

Determination of bias correction values

Correction values were calculated pooling data over sites, teeth and examiners, resulting in 4593 AL, 11155 PD and 5400 GH values. Each possible pairwise combination of periodontal probes (six in total) was considered. “Basic probe” designates the probe for which periodontal measurements need to be corrected to achieve comparability with periodontal measurements recorded with a second probe (designated as “target probe”). Mean values of target probes were evaluated for each level of the basic probe and used as correction values. The hierarchical structure of the data (examiner–subject–tooth–surface) was accounted for by application of STATA survey methods. As zero and high probing measurements were sparse ($N < 50$) and limit robust estimation of coefficients, zero and high values (i.e. AL ≥ 9 mm, PD ≥ 7 mm and GH $\geq 2/3$ mm) were retained, unchanged from the original observation.

Results

Periodontal examinations were performed on 2790 teeth (11160 sites). Using PCP11, PCP2 and PCPUNC15, AL was recorded at 1521, 1557 and 1515 sites respectively (Table 1). Corresponding numbers for PD were 3720, 3718 and 3717 sites respectively. GH was measured at 1813, 1774 and 1813 sites respectively. Patterns in missing measurements for AL, PD or GH across tooth type or measurement site were not associated with the probe ($p = 0.88$, chi-squared test, Table 1).

Distribution of measured periodontal values according to probe

For pairwise comparisons of distributions between periodontal probes (chi-squared test), AL ($p < 0.001$), PD ($p < 0.001$) and GH ($p < 0.001$) measurements differed significantly

between probes. Mean AL was 3.63 mm (95% CI, 1.52; 5.73), 3.41 mm (1.36; 5.47) and 3.50 mm (1.35; 5.65) for PCP11, PCP2 and PCPUNC15 respectively. Mean PD was highest for PCP11 (2.33 mm (1.39; 3.27)), followed by PCP2 (2.29 mm (1.33; 3.25)) and PCPUNC15 (2.25 mm (1.23; 3.28)). Mean GH was 1.33 mm (1.24; 1.41), 1.35 mm (1.26; 1.44) and 1.31 mm (1.21; 1.42) for PCP11, PCP2 and PCPUNC15 respectively. The probe-specific distribution of periodontal measurements was not restricted to certain examiners, but occurred globally (data not shown).

Values coinciding with probe graduation markings were selected more often by the examiners; an effect referred to as digit preference. For the PCP11 (Fig. 2), the AL (PD) value of 3 mm was named up to 1.4 (2.6) times more often than for the PCP2 or the PCPUNC15. For the PCP2, examiners reported an AL (PD) value of 2 mm up to 1.5 times more often than with one of the other two probes. For PD, the digit preference effect was also observed for higher thresholds (e.g. 4 mm), though to a minor degree. Higher probing values were too rare to observe any digit preference effects. For GH, a digit preference at 3 mm for PCP11 and at 2 mm for PCP2 was observable. For higher thresholds no assertions can be made.

Patterns of digit preference

Evaluating distance dimensions for two superimposed frequency distributions (Fig. 3), digit preference was most evident for PD measurements and less apparent for AL and GH measurements. PCPUNC15 was defined as gold standard.

Comparing distributions between PCP11 and PCPUNC15 (top figures), the digit preference at 3, 6 and 8 mm was most pronounced for PD (Fig. 3b). For PCP11, a PD value of 3 mm was reported about 12% more often compared with the PCPUNC15 probe. For AL and GH measurements (Fig. 3a,c), digit preference was still detectable, but less pronounced, especially at higher thresholds.

Values complying with the scale (i.e. 2, 4 and 6 mm) were reported more often using the PCP2 compared with the PCPUNC15 (bottom figures). A PD measurement of 2 mm was reported about 10% more often using the PCP2 compared with the PCPUNC15 (Fig. 3e). For AL and GH measurements (Fig. 3d,f), a digit preference was evident for the 2 mm marking.

Multilevel modelling of periodontal measurements

In multilevel modelling of periodontal measurements (Suppl. Table S1), the first carry-over effect and the interaction between probe and period were not statistically significant at the 10% level and were thus excluded from all models. For AL measurements, major effects for examiner ($p < 0.001$) and the interaction between probe and examiner were found ($p < 0.001$). Main effects for period and probe were non-significant. The difference between PCPUNC15 and PCP2 was not statistically significant ($p = 0.58$). For PD measurements, a statistically significant effect of probe ($p < 0.001$), accompanied by a major effect of examiner ($p < 0.001$) and an interaction between probe and examiner ($p < 0.001$), was revealed. All three probes differed significantly among each other ($p < 0.05$). Evaluating GH measurements, major effects were found for probe ($p < 0.001$), examiner ($p < 0.001$) and the interaction between probe and examiner ($p < 0.001$). PCP2 and PCPUNC15 differed significantly from PCP11 ($p = 0.001$ and $p < 0.001$ respectively). The difference between PCPUNC15 and PCP2 missed statistical significance ($p = 0.06$).

Bias correction of AL, PD and GH measurements

Correction values for AL, PD and GH measurements are listed in Table 2. To illustrate, if PCP11 was the basic probe and PCP2 was the target probe, AL values of 1–8 mm would be replaced by correction values of 1.39, 2.06, 2.77, 3.84, 4.81, 5.65, 6.68 and 7.47 mm respectively. AL values of 0 and 9 mm would be retained. Accordingly, PD values of 1–6 mm would be replaced by 1.35, 1.95, 2.65, 3.80, 4.63 and 5.73 mm respectively. PD values of 0 and 7 mm would be retained. GH values of 1–3 mm would be replaced by 1.12, 1.86 and 2.09 mm respectively. GH values of 0 and 4 mm would be retained. Based on the corrected site measurements, mean AL, mean PD and mean GH on subject level would be calculated for further analyses.

The application of the proposed correction values for PD measurement assessed with the PCP11 (basic probe) and the PCP2 (target probe) in two hypothetical studies is given in Fig. 4.

Discussion

This full factorial crossover design enabled direct comparison of three manual periodontal probes in a rigorous study design. In this crossover study, distributions of periodontal measures differed between the three periodontal probes, PCP11, PCP2 and PCPUNC15. For PCP11 and PCP2, we observed a distinct accumulation of values complying with the respective probe graduation; an effect referred to as digit preference. The PCPUNC15 was less affected. Digit preference was most pronounced for PD measurements and less pronounced, but still detectable, for AL and GH measurements. Correction values for pairwise combinations of probes were derived to enable comparison of periodontal distributions on subject level between studies with identical recording protocols, but different periodontal probes.

According to multilevel modeling, periodontal measurements were mainly affected by examiner and the periodontal probe (Suppl. Table S1). Differences in periodontal probing due to examiner might be explained by differences in experience and calibration quality. Despite training of examiners, variations in probing forces and probing technique need to be assumed (Hassel et al. 1973). However, as similarity of examiners between different surveys, even within longitudinal studies, cannot be assumed, correction of probe measurements for examiner differences is not feasible. Thus, probe differences are the only actionable element to reduce variability.

Several factors might contribute to probe differences in measuring periodontal parameters. In our study, on each examination day, the sequence of periodontal probes for each subject was randomly assigned. Thus, an effect of bleeding on probing, increased indisposition of study subjects or decreasing concentration of examiners on penetration depth is unlikely. As all handles had equal diameters, effects of handle diameter on probing pressure (van Weringh et al. 2006) can be excluded. Relevant probe-related reasons for measurement differences between the three probes might be found (i) in the corrugated structure of the probe surface, (ii) in the diameter and shape of the probe tip (Atassi et al. 1992, Garnick & Silverstein 2000) and (iii) in the kind of graduation markings.

Probe surfaces differed for PCPUNC15 as compared with PCP11 and PCP2. The PCPUNC15 probe has a 1 mm scale with coloured, engraved and corrugated graduation markings, whereas the PCP11 and the PCP2 have coloured and smooth graduations. Possibly, probes with a rougher surface might not penetrate as deep into periodontal pockets as smoother ones, assuming identical penetration forces. Thus, a slight underestimation of periodontal pockets for the PCPUNC15 might be assumed.

The diameter and shape of the probe might influence penetration depths (Van der Zee et al. 1991, Bulthuis et al. 1998 and Garnick & Silverstein 2000). Probes with thick tip diameters cannot penetrate as deep as probes with smaller tip diameters (Van der Zee et al. 1991). Especially after resolution of inflammation, when initial periodontal treatment has resulted in formation of a long junctional epithelium, thick probes might underestimate pocket depth even with relatively high probing pressures (Van der Zee et al. 1991). In contrast, using thin probes might result in excessive probing (Garnick & Silverstein 2000); minimal probing pressures are sufficient to penetrate deep-inflamed pockets to deeper depths (Bulthuis et al. 1998, Barendregt et al. 2008). In this study, the tine diameter was highest for PCPUNC15 (0.57 mm) – almost equalling the suggested optimal diameter of 0.6 mm by Garnick & Silverstein (2000). Tip diameters for PCP11 and PCP2 were 0.48 and 0.50 mm respectively. Assuming a low level of tissue inflammation in studied subjects, results concur with the literature: on average, the thickest probe (PCPUNC15) resulted in lower PD measurements than PCP11 and PCP2. However, assuming that thin probes (0.1–0.5 mm) “over-probed the sulcus and extended to the base of the junctional epithelium” (Garnick & Silverstein 2000), PCP11 and PCP2 might have slightly overestimated pocket depth.

Finally, digit preference might provide a plausible explanation for probe-related differences in periodontal measurements. The digit preference effect was well documented in medical literature (Rowland 1990, Crawford et al. 2002, Thavarajah et al. 2003, Camarda et al. 2008). Digit preference describes rounding bias associated with specific instruments and can bias the observed distribution of values (Camarda et al. 2008). In this study, frequency distributions of periodontal measurements perfectly agreed with respective graduations of all three manual probes, especially for slight and moderate values, that is 2, 3 and 4 mm. As most subjects in epidemiological studies exhibit slight to moderate PD and AL values (Holtfreter et al. 2009), this effect is of major importance. Especially in longitudinal studies, use of different periodontal probes may result in biased estimates of periodontal progression.

Correction values (Table 2) provide a possibility to alleviate probe-associated bias in case of identical recording protocols. In practical, periodontal site measurements are replaced by correction values of the target probe and mean values for AL, PD or GH over all assessed measurement sites in one subject are calculated. However, use of correction values is limited to population-wide analyses using subject-level averaged values. Correction values may not be applied if alternative periodontal measures like dichotomization by using cut-off values are used (e.g. relating to the percentage of sites exhibiting fixed cut-offs, e.g. 3 mm). As the measurements in question are shifted to higher values, for example, from 3 (PCP2) to 3.11 mm (PCP11), the extent of PD $\geq$ 3 mm would not change. If measurements were shifted to lower values, for example, from 3 (PCP2) to 2.94 mm (PCPUNC15), estimates of extent of PD $\geq$ 3 mm after probe correction would equal the extent of PD $\geq$ 4 mm before probe correction. This turns out to be problematic, because after probe correction, the percentage of sites exhibiting PD $\geq$ 3 mm is reduced by the percentage of sites with PD = 3 mm before probe correction.

Some limitations deserve consideration. We cannot provide reliable correction values for zero and high measurements, that is PD $\geq$ 7 mm, AL $\geq$ 9 mm and GH $\geq$ 2/3 mm, because of too sparse data limiting robust estimation of coefficients. Thus, zero and high values were left unchanged from the original observation. Future studies should recruit more subjects showing a broader spectrum of periodontal disease severity. Further limitations connected with this study design comprehend short intervals between measurements within each day and missing information on clinical status before the study. Intervals between examinations were kept short (1st day: median 8 min, range 0–17) to avoid drop outs among study participants and to reduce overall time and costs. To which extent repeated measurements might have led to increasing probing depths (Osborn et al. 1990) cannot be assessed using

the present data. Importantly, the order of probes and examiner within each subject was random. Thus, bleeding on probing or deepened pockets due to excessive probing should not be systematically related to probe type or examiner.

To conclude, compared with the PCP11 and the PCP2, the PCPUNC15 might provide more exact periodontal measurements assuming proper examiner training and calibration. Thus, the PCPUNC15 should preferentially be used in clinical and epidemiological studies. As of its fine and equidistant scaling, it allows for higher accuracy (Van der Zee et al. 1991) and thus facilitates correct mathematical rounding of periodontal measurements, which is of special importance (Tibbetts 1969). This fine scaling also limits the opportunity for digit preference. In terms of reliability and accuracy, manual probes with 1 mm scaling ranked well among automated and conventional probes (Van der Zee et al. 1991, Grossi et al. 1996, Samuel et al. 1997, Karpinia et al. 2004) with experience having only a slight effect on reproducibility (Samuel et al. 1997). Further, Van der Zee et al. (1991) found that probes with scribed bands, representing marking lines of no appreciable width, were optimal and had highest accuracy compared with etched markings, engraved grooves and painted bands.

This study has further implications. As many studies used the WHO probe to estimate periodontal treatment needs (Leroy et al. 2010), the impact and direction of digit preference bias (at 3.5, 5.5, 8.5 or 11.5 mm markings, as reported here) on periodontal treatment needs should be studied. Assuming a comparable digit preference as for the PCP11, which has a similar graduation, severe over-estimation of moderate and severe periodontal disease might be expected. The importance of maintaining the same probe in longitudinal population studies can also be inferred from our findings.

In summary, we provided empirical evidence and quantified the impact of probes on periodontal measurements in a rigorously designed crossover study of the most commonly used instruments. Differences in periodontal measure distributions between PCP11 and PCP2 as compared with the PCPUNC15 might well be explained by effects of digit preference seen for both probes. We also provided correction factors that may be utilized to optimize comparison across clinical and epidemiological studies, assuming data access. In contrast to PCPUNC15, which is proposed as a gold standard for manual periodontal probing, the PCP11 and PCP2 might be more suitable for screening purposes. Importantly, the standardization of periodontal probes to enhance the accuracy and reproducibility of periodontal measurements in clinical and epidemiological studies should be considered.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Source of funding:

Study of Health in Pomerania (SHIP) is a part of the Community Medicine Net (<http://www.medizin.uni-greifswald.de/cm>) of the University of Greifswald, which is funded by grants from the German Federal Ministry of Education and Research (BMBF, grants no. 01ZZ9603, 01ZZ0103 and 01ZZ0403); the Ministry for Education, Research, and Cultural Affairs; and the Ministry for Social Affairs of the Federal State of Mecklenburg- West Pomerania. B.H. was supported by an unlimited educational grant from GABA, Switzerland. M.D. was financed by an INSERM Chair of Excellence and the Chair in Chronic Disease from the École des Hautes Études en Santé Publique.

References

- Atassi F, Newman HN, Bulman JS. Probe tine diameter and probing depth. *Journal of Clinical Periodontology*. 1992; 19:301–304. [PubMed: 1517473]
- Barendregt DS, van der Velden U, Timmerman MF, van der Weijden F. Penetration depths with an ultrasonic mini insert compared with a conventional curette in patients with periodontitis and in periodontal maintenance. *Journal of Clinical Periodontology*. 2008; 35:31–36. [PubMed: 18173399]
- Beck JD, Caplan DJ, Preisser JS, Moss K. Reducing the bias of probing depth and attachment level estimates using random partial-mouth recording. *Community Dentistry and Oral Epidemiology*. 2006; 34:1–10. [PubMed: 16423025]
- Bulthuis HM, Barendregt DS, Timmerman MF, Loos BG, van der Velden U. Probe penetration in relation to the connective tissue attachment level: influence of tine shape and probing force. *Journal of Clinical Periodontology*. 1998; 25:417–423. [PubMed: 9650880]
- Camarda GC, Eilers PHC, Gampe J. Modelling general patterns of digit preference. *Statistical Modelling*. 2008; 8:385–401.
- Crawford SL, Johannes CB, Stellato RK. Assessment of digit preference in self-reported year at menopause: choice of an appropriate reference distribution. *American Journal of Epidemiology*. 2002; 156:676–683. [PubMed: 12244037]
- Fleiss JL, Mann J, Paik M, Goultschin J, Chilton NW. A study of inter- and intra-examiner reliability of pocket depth and attachment level. *Journal of Periodontal Research*. 1991; 26:122–128. [PubMed: 1826526]
- Garnick JJ, Silverstein L. Periodontal probing: probe tip diameter. *Journal of Periodontology*. 2000; 71:96–103. [PubMed: 10695944]
- Grossi SG, Dunford RG, Ho A, Koch G, Machtei EE, Genco RJ. Sources of error for periodontal probing measurements. *Journal of Periodontal Research*. 1996; 31:330–336. [PubMed: 8858537]
- Hassel TM, Germann MA, Saxer UP. Periodontal probing: interinvestigator discrepancies and correlations between probing force and recorded depth. *Helvetica Odontologica Acta*. 1973; 17:38–42. [PubMed: 4699376]
- Hefti AF. Periodontal probing. *Critical Reviews in Oral Biology and Medicine*. 1997; 8:336–356. [PubMed: 9260047]
- Hensel E, Gesch D, Biffar R, Bernhardt O, Kocher T, Splieth C, Born G, John U. Study of Health in Pomerania (SHIP): a health survey in an East German region. Objectives and design of the oral health section. *Quintessence International*. 2003; 34:370–378. [PubMed: 12795357]
- Holtfreter B, Schwahn C, Biffar R, Kocher T. Epidemiology of periodontal diseases in the Study of Health in Pomerania. *Journal of Clinical Periodontology*. 2009; 36:114–123. [PubMed: 19207886]
- Karpinia K, Magnusson I, Gibbs C, Yang MC. Accuracy of probing attachment levels using a CEJ probe versus traditional probes. *Journal of Clinical Periodontology*. 2004; 31:173–176. [PubMed: 15016020]
- Kashmere JL, Kirk A. The complex interaction of normal biases in line bisection. *Neurology*. 1997; 49:887–889. [PubMed: 9305363]
- Larsen C, Barendregt DS, Slot DE, Van der Velden U, Van der Weijden F. Probing pressure, a highly undervalued unit of measure in periodontal probing: a systematic review on its effect on probing pocket depth. *Journal of Clinical Periodontology*. 2009; 36:315–322. [PubMed: 19426178]
- Leroy R, Eaton KA, Savage A. Methodological issues in epidemiological studies of periodontitis—how can it be improved? *BMC Oral Health*. 2010; 10:8. [PubMed: 20409298]
- Listgarten MA. Periodontal probing: what does it mean? *Journal of Clinical Periodontology*. 1980; 7:165–176. [PubMed: 7000852]
- Osborn J, Stollenberg J, Huso B, Aepli D, Pihlstrom B. Comparison of measurement variability using a standard and constant force periodontal probe. *Journal of Periodontology*. 1990; 61:497–503. [PubMed: 2391627]
- Persson GR. Effects of line-angle versus mid-proximal periodontal probing measurements on prevalence estimates of periodontal disease. *Journal of Periodontal Research*. 1991; 26:527–529. [PubMed: 1837057]

- R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2011.
- Reddy MS, Palcanis KG, Geurs NC. A comparison of manual and controlled-force attachment-level measurements. *Journal of Clinical Periodontology*. 1997; 24:920–926. [PubMed: 9442430]
- Rowland ML. Self-reported weight and height. *American Journal of Clinical Nutrition*. 1990; 52:1125–1133. [PubMed: 2239790]
- Samuel ED, Griffiths GS, Petrie A. In vitro accuracy and reproducibility of automated and conventional periodontal probes. *Journal of Clinical Periodontology*. 1997; 24:340–345. [PubMed: 9178114]
- Silva-Boghossian CM, Amaral CS, Maia LC, Luiz RR, Colombo AP. Manual and electronic probing of the periodontal attachment level in untreated periodontitis: a systematic review. *Journal of Dentistry*. 2008; 36:651–657. [PubMed: 18534736]
- StataCorp. Stata Statistical Software: Release 10. College Station, TX: StataCorp LP; 2007.
- Susin C, Kingman A, Albandar JM. Effect of partial recording protocols on estimates of prevalence of periodontal disease. *Journal of Periodontology*. 2005; 76:262–267. [PubMed: 15974851]
- Thavarajah S, White WB, Mansoor GA. Terminal digit bias in a specialty hypertension faculty practice. *Journal of Human Hypertension*. 2003; 17:819–822. [PubMed: 14704725]
- Tibbetts LS. Use of diagnostic probes for detection of periodontal disease. *Journal of the American Dental Association*. 1969; 78:549–555. [PubMed: 5249926]
- Van der Zee E, Davies EH, Newman HN. Marking width, calibration from tip and tine diameter of periodontal probes. *Journal of Clinical Periodontology*. 1991; 18:516–520. [PubMed: 1894745]
- Vettore MV, de Lamarca GA, Leao AT, Sheiham A, do Leal MC. Partial recording protocols for periodontal disease assessment in epidemiological surveys. *Cad Saude Publica*. 2007; 23:33–42. [PubMed: 17187102]
- Volzke H, Alte D, Schmidt CO, Radke D, Lörber R, Friedrich N, Aumann N, Lau K, Piontek M, Born G, Havemann C, Ittermann T, Schipf S, Haring R, Baumeister SE, Wallaschofski H, Nauck M, Frick S, Arnold A, Junger M, Mayerle J, Kraft M, Lerch MM, Dorr M, Reffelmann T, Empen K, Felix SB, Obst A, Koch B, Glaser S, Ewert R, Fietze I, Penzel T, Doren M, Rathmann W, Haerting J, Hannemann M, Ropcke J, Schminke U, Jurgens C, Tost F, Rettig R, Kors JA, Ungerer S, Hegenscheid K, Kuhn JP, Kuhn J, Hosten N, Puls R, Henke J, Gloger O, Teumer A, Homuth G, Volker U, Schwahn C, Holtfreter B, Polzer I, Kohlmann T, Grabe HJ, Roskopf D, Kroemer HK, Kocher T, Biffar R, John U, Hoffmann W. Cohort profile: the study of health in Pomerania. *International Journal of Epidemiology*. 2011; 40:294–307. [PubMed: 20167617]
- Watts T. Constant force probing with and without a stent in untreated periodontal disease: the clinical reproducibility problem and possible sources of error. *Journal of Clinical Periodontology*. 1987; 14:407–411. [PubMed: 3476518]
- van Weringh M, Barendregt DS, Rosema NA, Timmerman MF, van der Weijden GA. A thin or thick probe handle: does it make a difference? *International Journal of Dental Hygiene*. 2006; 4:140–144. [PubMed: 16958742]

Clinical Relevance

Scientific rationale for the study

Differences in periodontal measurements were assessed for three manual periodontal probes with different graduation.

Principal findings

Measurements of attachment loss, probing depth and gingiva height differed significantly between PCP2/PCPUNC15 and PCP11. The digit preference effect provided a plausible explanation. Correction values were determined to enable comparison between studies with identical recording protocols, but different periodontal probes.

Practical implications

This study emphasizes the importance of consistency in the periodontal probe type for clinical and epidemiological studies. PCP11 and PCP2 are recommended for screening purposes, whereas PCPUNC15 might provide exact measurements assuming proper examiner training and calibration.

Fig. 1. Micrographs for (a,d) PCP11, (b,e) PCP2 and (c,f) PCPUNC15 with 5-fold (top figures) and 25-fold magnification (bottom figures).

Fig. 2. Frequency distribution of attachment loss (left column), probing depth (middle column), and gingiva height (right column) for three manual periodontal probes (colouring).

Fig. 3.

Difference between relative frequency distributions for pairwise combinations of periodontal probes (PCP11 and PCP2) in comparison with PCPUNC15 for attachment loss (a,d), probing depth (b,e) and gingiva height (c,f). Vertical lines indicate engraved markings for PCP11 (above) or PCP2 (below).

Fig. 4. Hypothetical work flow for the application of correction values for probing depth measurements assessed with the PCP11 (Study A, basic probe) and the PCP2 (Study B, target probe).

Table 1
Characteristics of examined teeth and sites pooled over all examiners and periods. Attachment loss Probing depth Gingiva height

Location	Attachment loss				Probing depth				Gingiva height			
	PCP11 No. (%)	PCP2 No. (%)	PCPUNC15 No. (%)	Total	PCP11 No. (%)	PCP2 No. (%)	PCPUNC15 No. (%)	Total	PCP11 No. (%)	PCP2 No. (%)	PCPUNC15 No. (%)	Total
Anterior teeth	701 (46.1)	712 (45.7)	682 (45.0)	2095	1632 (43.9)	1630 (43.8)	1631 (43.9)	4893	809 (44.6)	795 (44.8)	823 (45.4)	2427
Premolar	440 (28.9)	448 (28.8)	448 (29.6)	1336	1128 (30.3)	1128 (30.3)	1128 (30.3)	3384	520 (28.7)	512 (28.9)	511 (28.2)	1543
Molar	380 (25.0)	397 (25.5)	385 (25.4)	1162	960 (25.8)	960 (25.9)	958 (25.8)	2878	484 (26.7)	467 (26.3)	479 (26.4)	1430
Site				<i>p</i> = 0.98 *				<i>p</i> = 1.00 *				<i>p</i> = 0.99 *
Distobuccal	346 (22.7)	344 (22.1)	328 (21.7)	1018	931 (25.0)	930 (25.0)	929 (25.0)	1925	488 (27.0)	490 (27.6)	502 (27.7)	1480
Midbuccal	482 (31.7)	518 (33.3)	489 (32.3)	1489	930 (25.0)	930 (25.0)	929 (25.0)	2789	352 (19.4)	316 (17.8)	344 (19.0)	1012
Mesiobuccal	285 (18.7)	273 (17.5)	287 (18.9)	845	930 (25.0)	930 (25.0)	930 (25.0)	2790	548 (30.2)	560 (31.6)	545 (30.1)	1653
Midlingual/midpalatal	408 (26.8)	422 (27.1)	411 (27.1)	1241	929 (25.0)	928 (25.0)	929 (25.0)	2788	425 (23.4)	408 (23.0)	422 (23.2)	1255
Total	1521 (100)	1557 (100)	1515 (100)	4593	3720 (100)	3718 (100)	3717 (100)	11,155	1813 (100)	1774 (100)	1813 (100)	5400
				<i>p</i> = 0.92 *				<i>p</i> = 1.00 *				<i>p</i> = 0.88 *

Data are presented as numbers (percentages).

* Chi-squared test.

Comparison of periodontal measurements for pairwise combinations of periodontal probes. “Basic probe” designates the probe for which periodontal measurements need to be corrected to achieve comparability with periodontal measurements recorded with a second probe (designated as “target probe”). There are three blocks corresponding to three possible basic probes (marked bold): PCP11, PCP2 and PCPUNC15. For each of the three blocks, millimeter levels of the basic probe, the number (*N*) of measurement sites exhibiting respective millimeter values using the basic probe and the correction value are given

Table 2

PCP11...		...to PCP2		...to PCPUNC15		PCP2...		...to PCP11		...to PCPUNC15...		...to PCP2		...to PCP11	
mm level	N	Correction value	N	Correction value	N	mm level	mm level	N	Correction value	N	Correction value	N	Correction value	N	Correction value
Attachment loss, mm															
1	161	1.39	140	1.39	139	1	1	139	1.33	133	1.34	154	1.47	151	1.52
2	233	2.06	237	2.11	347	2	2	347	2.29	350	2.18	300	2.12	279	2.29
3	275	2.77	269	2.81	207	3	3	207	3.11	207	2.94	235	2.94	231	3.06
4	169	3.84	168	3.92	157	4	4	157	4.09	154	4.05	171	4.24	176	4.19
5	126	4.81	124	4.85	135	5	5	135	4.96	126	5.06	106	4.91	113	5.04
6	111	5.65	111	6.04	113	6	6	113	6.27	113	6.09	107	5.77	113	6.16
7	92	6.68	94	6.57	76	7	7	76	6.89	72	6.42	87	6.46	93	6.66
8	58	7.47	61	6.84	51	8	8	51	7.35	51	7.16	49	7.49	52	7.52
Probing depth, mm															
1	1064	1.35	1064	1.25	874	1	1	874	1.21	874	1.17	1123	1.37	1123	1.32
2	1290	1.95	1291	1.92	1941	2	2	1941	2.11	1941	2.00	1560	2.04	1561	2.14
3	982	2.65	980	2.69	372	3	3	372	2.86	371	2.77	535	2.72	535	2.85
4	102	3.80	102	4.00	277	4	4	277	3.57	276	3.75	192	3.70	192	3.63
5	93	4.63	93	4.56	88	5	5	88	5.05	88	4.83	162	4.44	162	4.51
6	131	5.73	131	5.50	112	6	6	112	5.90	112	5.55	95	5.75	95	5.85
Gingiva height, mm															
1	1012	1.12	1030	1.11	968	1	1	968	1.08	939	1.10	987	1.15	1016	1.10
2	430	1.86	431	1.81	521	2	2	521	1.87	538	1.78	484	1.82	486	1.86
3	68	2.09	67	2.22											

Bias correction values equal the mean of target probe measurements evaluated for each millimeter level of the basic probe. Correlation within subjects was accounted for via survey methods provided in STATA.