Accuracy of Partial-Mouth Examination Protocols for Extent and Severity Estimates of Periodontitis: A Study in a Chinese Population with Chronic Periodontitis

Journal:	*Journal of Periodontology*
Manuscript ID:	JOP-14-0422.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	04-Oct-2014
Complete List of Authors:	Chu, Yi; Peking University School and Hospital of Stomatology, Department of Periodontology Ouyang, Xiangying; Peking University School and Hospital of Stomatology, Department of Periodontology
Key Words:	Epidemiology, Periodontitis, Diagnosis

Accuracy of Partial-Mouth Examination Protocols for Extent and Severity Estimates of Periodontitis: A Study in a Chinese Population with Chronic Periodontitis

Yi Chu DDS[*†], Xiangying Ouyang DDS[*]

[*] Department of Periodontology, School and Hospital of Stomatology, Peking University, Beijing, China.

[†] Department of Outpatient Center, School and Hospital of Stomatology, Peking University, Beijing, China.

Address for Correspondence: Prof. Ouyang Xiangying, Department of Periodontology, Peking University School and Hospital of Stomatology, 22 Zhongguancun Nandajie, Haidian District, 100081, Beijing, China.

Fax number: 86-10-62173402

E-mail: kqouyangxy@bjmu.edu.cn

E-mail can be published and fax number better not be published.

Word count: Abstract: 253; Text: 3928.

Number of Figures and Tables: Four tables and 2 figures in the print publication, and 3 figures in supplement files online.

Running title: Partial-mouth Protocols: Accuracy in Extent and Severity

One-sentence summary: Using data from a Chinese population with periodontitis, random half-mouth six-sites per tooth protocol produced the least biased and most consistent estimates for the extent and severity of attachment loss, probing depth, and bleeding on probing of the fifteen partial-mouth examination protocols investigated.

ABSTRACT

Background Partial-mouth periodontal examination (PMPE) has been widely used in periodontal epidemiological studies. In this study, we evaluated the accuracy of extent and severity estimates from PMPE protocols in a Chinese population.

Methods We enrolled 200 individuals with periodontitis, aged 22 to 64 years. Full-mouth examination was performed to determine probing depth (PD), attachment loss (AL), and bleeding on probing (BOP) at mesial-buccal (MB), mid-buccal (B), disto-buccal (DB), mesial-lingual (ML), mid-lingual (L), and disto-lingual (DL) sites per tooth. Extent and severity estimates from fifteen PMPE protocols were derived from and compared to full-mouth data. Relative bias (RB) and intra-class correlation coefficient (ICCs) were calculated. Bland-Altman plots were used to evaluate the agreement patterns across disease levels.

Results Of the fifteen PMPE protocols, random half-mouth six-sites per tooth (r6sites) protocol performed best in both extent (AL ≥ 2, 4 or 6 mm; PD ≥ 4 or 6 mm; and BOP) and severity (AL and PD) estimates, with RB within ±5.0% and ICCs ≥0.950 in most cases. MB-B-DB and MB-B-DL protocols generally caused RB within ±20.0% for extent and within ±5.0% for severity. Protocols only involving interproximal sites (MB-DB, MB-DL, and MB-DB-ML-DL) showed good accuracy in AL (RB within ±20.0% for extent; within ±3.0% for severity), but overestimated PD (RB 12.5%-54.2% for extent; >10.0% for severity). The community periodontal index teeth protocol caused severe overestimation of up to 110.4% for extent and 14.6% for severity.

Conclusions The r6sites protocol is best for assessing extent and severity for AL, PD, and BOP under the study conditions.

KEY WORDS Health surveys, periodontitis, bias, periodontal index

Introduction

In epidemiological studies of periodontal disease, full-mouth periodontal examination (FMPE) is currently used as the gold-standard protocol[1]. This method examines six sites per tooth: the mesio-buccal (MB), mid-buccal (B), disto-buccal (DB), mesio-lingual (ML), mid-lingual (L), and disto-lingual (DL) sites[2]. FMPE requires the probing of a maximum of 168 sites (if third molars are excluded), and examination time can range from 28.8[3] to 40 minutes[4]. It has been time-consuming to use FMPE in large-scale periodontal surveys. Moreover, probing full-mouth sites may cause fatigue for both the examiner and participant, which is likely to cause an increase in measurement errors[5]. Based on symmetry and site- and tooth-specific susceptibility to periodontal diseases, various partial-mouth periodontal examination (PMPE) protocols have been reported and used for most large-scale epidemiological studies[3, 6-14].

PMPE protocols can be classified into two types: index-teeth protocols and non-index-teeth protocols. The most widely-used index-teeth are Ramfjord teeth[6] and community periodontal index (CPI) teeth[7]. Non-index-teeth protocols are usually based on a full- or half-mouth design that examines ≤6 sites per tooth[10, 11, 13-15], such as the random half-mouth MB-B protocol used in the National Health and Nutrition Examination Survey (NHANES) III from 1999–2000[10] and the random half-mouth MB-B-DB protocol used in the NHANES from 2001–2004[15]. PMPE protocols that examined fewer sites, teeth or quadrants reduced examination time, but also caused biases in disease estimates[2-5, 12, 14, 16-35].

That PMPE underestimated prevalence levels when compared with FMPE has been well demonstrated[2-4, 12, 14, 16-25, 27-32, 34, 35], and the accuracy of PMPE prevalence estimates improved as the number of examined sites increased[34]. However, the accuracy of extent and severity estimates

from the various PMPE protocols has been less clear. Previous studies have assessed the reliability of PMPE protocols for extent and severity estimates[2, 3, 5, 12, 17, 19-22, 24-26, 29, 35]. Some used protocols that evaluated fewer than six sites per tooth as the "gold-standard" [17, 19-22], while some analyzed severity estimates alone[2, 5, 25, 26]. However, severity determined by mean clinical values may have led to data flattening, because sites exhibiting more severe disease tended to be obscured by unaffected sites. Extent is expressed as the percentage of sites that fall above a specific severity threshold. Therefore, assessing both extent and severity simultaneously more fully describes the disease level of a population.

Fewer studies have evaluated the accuracy of extent and severity at the same time, and only a limited number of PMPE protocols were investigated in each study, including the index teeth[3, 12, 24, 29, 35], MB-B[3, 12], MB-B-DB[12, 29], and half-mouth six-sites per tooth methods[3, 24, 29, 35]. The MB-B-DL protocols, which performed well in prevalence and severity estimates[5, 27], had not been previously evaluated for extent scores. A very recent study evaluated interproximal sites protocols such as MB-DB, MB-DL and MB-DB-ML-DL in prevalence scores[14]. However, their performance in extent and severity estimates has been less clear. Moreover, since these protocols were assessed in different populations, it has been difficult to compare data across these studies to determine the most accurate PMPE protocols. These protocols have not been compared thoroughly in a single sample population, so no consensus has been built concerning which PMPE protocol had best accuracy for extent and severity estimates[34].

In addition, the accuracy of a PMPE protocol might be associated with the demographic characteristics and disease level of the population of interest[5]. A recent systemic review evaluated the performance of PMPE protocols in probing depth (PD) / attachment loss (AL) ≥4mm

thresholds[34]. Evaluations of more severe periodontal disease thresholds have been lacking[3, 34]. Moreover, most previous analyses were conducted in American[2, 3, 5, 12, 24, 25, 29] , European[35] or African[26] populations. To the best of our knowledge, no assessments have been performed on Chinese populations.

Therefore, the aim of this study was to assess the accuracy of extent and severity estimates from PMPE protocols in a Chinese population with periodontitis.

Materials and methods

The dataset from this study came from the data of a cross-sectional epidemiological study for chronic periodontitis performed in the year 2000 in Beijing. Detailed information concerning the methodological procedures of the 2000 survey has been published elsewhere[36]. Briefly, five hundred individuals working in an institute were screened. Two hundred individuals with periodontitis were selected as the study sample and took part in a full-mouth examination. The inclusion criteria were that all individuals should have at least 16 teeth, at least four molars, a ≥5 mm PD and ≥2 mm AL in at least two sites in different quadrants[10]. No participants received any periodontal therapy or antibiotics within the previous year. Participants were excluded if they were pregnant or had systemic conditions that required prophylactic antibiotics prior to periodontal examination. Of the 200 participants examined, 112 (56.0%) were male, and 55 (37.5%) were current smokers. Ages ranged from 22 to 64 years with a mean age of 43.1 ± 10.8 years.

This survey performed in 2000 was approved by the institutional review board of Procter & Gamble Corporation, and written informed consent was obtained from each subject in accordance

with the Declaration of Helsinki. We re-analyzed the data of this 2000 survey for accuracy of partial-mouth examination protocols, and this re-analysis study was approved by the institutional review board of Peking University School and Hospital of Stomatology (PKUSSIRB-201413037).

Periodontal examination

All of the participants were clinically examined in the institute equipped with a complete dental unit in the field. Clinical examination included PD and AL of all erupted permanent teeth, excluding third molars. Each tooth was examined at six sites: MB, B, DB, ML, L, and DL. A manual periodontal probe was used‡. Measurements were made in millimeters and rounded to the nearest whole millimeter. Thirty seconds after probing, the presence or absence of bleeding at probed sites was recorded.

All examinations were conducted by a single experienced examiner (OX). To demonstrate measurement reproducibility, ten subjects from the sample were randomly selected and examined again. Full-mouth examination showed high reproducibility for PD, AL and BOP. The intra-class correlation coefficient (ICC) for PD was 0.909. The ICC for AL was 0.884, and the κ value for BOP was 0.83.

Partial-mouth examination protocols

All partial-mouth datasets were derived from full-mouth data. In this study, evaluation was conducted in fifteen PMPE protocols for comparison with the FMPE results. These protocols

were random half-mouth six-sites per tooth (r6sites); full-mouth MB-B-DB (fMB-B-DB); random half-mouth MB-B-DB (rMB-B-DB); full-mouth MB-B-DL (fMB-B-DL); random half-mouth MB-B-DL (rMB-B-DL); full-mouth MB-B (fMB-B); random half-mouth MB-B (rMB-B); full-mouth MB-DB (fMB-DB); random half-mouth MB-DB (rMB-DB); full-mouth MB-DL (fMB-DL); random half-mouth MB-DL (rMB-DL); full-mouth MB-DB-ML-DL (fMB-DB-ML-DL); random half-mouth MB-DB-ML-DL (rMB-DB-ML-DL); Ramfjord teeth six-sites per tooth (RT) ; and CPI teeth six-sites per tooth (CPI). The Ramfjord teeth were #3, #9, #12, #19, #25, and #28, and the CPI teeth were #2, #3, #8, #14, #15, #18, #19, #24, #30, and #31. In random half-mouth protocols, diagonal quadrants were randomly selected for each subject.

Data analysis

The extent score for a specific individual was determined as the percentage of sites above a certain threshold. For the whole sample, extent estimates were the mean extent scores for each participant. The severity score for a particular individual was calculated from mean measurements of each site. For the whole sample, severity estimates were calculated from the mean severity scores for each participant. Extent estimates for AL ≥2, 4 or 6 mm and PD ≥4 or 6 mm; BOP; and severity estimates for AL and PD were calculated using data from the FMPE protocol and each PMPE protocol.

To assess the systemic biases of each PMPE protocol for the whole sample, relative biases was calculated[2] as the:

relative bias of extent estimates for the whole sample = $(E_{PMPE}-E_{FMPE}) / E_{FMPE} \times 100\%$; and

relative bias of severity estimates for the whole sample = $(S_{PMPE}-S_{FMPE}) / S_{FMPE} \times 100\%$;

where:

E_{PMPE} = sample extent estimates determined by the PMPE protocol;

E_{FMPE} = sample extent estimates determined by the FMPE protocol;

S_{PMPE} = sample severity estimates determined by the PMPE protocol; and

S_{FMPE} = sample severity estimates determined by the FMPE protocol.

Relative biases reflected the percentage of over/underestimation compared to the FMPE protocol. A positive relative bias indicated that PMPE protocols overestimated the results of the FMPE protocol, and a negative relative bias indicated underestimation. Relative biases were calculated for extent of AL ≥2, 4 or 6 mm and PD≥ 4 or 6mm; BOP; and the severity of AL and PD for each PMPE protocol. Differences between each PMPE protocol and FMPE protocol were tested with the paired t-test or the Wilcoxon-signed rank test§.

To assess the agreement between PMPE and FMPE scores, intra-class correlation coefficients (ICCs) were calculated§ for extent of AL ≥2, 4 or 6 mm; PD ≥4 or 6 mm; BOP; and the severity of AL and PD, according to the methodology described by McGraw[37] and Kingman et al[5]. In addition, Bland-Altman plots[38] were used to investigate the agreement pattern across a range of disease levels by plotting differences in PMPE and FMPE individual scores against their average scores‖. For each PMPE protocol, extent of AL ≥4 mm, PD ≥4 mm and BOP, and severity of AL and PD were analyzed with Bland-Altman plots.

Results

Fifty-seven percent of the participants had 28 teeth (excluding third molars), 42.5% had less than

five lost teeth, and 0.5% had more than six lost teeth. All the participants had at least one AL ≥2 mm site, and AL ≥4 mm sites were found in 87.0% of all participants. The prevalence of AL ≥6 mm was 52.0%. All participants had at least one site with PD ≥4 mm, and a PD ≥6 mm site affected 96.0% of the sample. All participants had BOP positive sites.

Bias for extent estimates

Table 1 and 2 show the extent estimates of the sample and relative biases for each PMPE protocol. The relative biases were negative or positive, indicating either under- or overestimation, respectively. The absolute value of the relative bias tended to increase as the severity thresholds increased. When the same type of sites were examined, the full- and random half-mouth protocols showed similar sample extent estimates, even though the numbers of examined sites were different. For example, the sample extent estimate for AL ≥2 mm was 49.4% and 49.9% in fMB-B-DB (84 sites per mouth) and rMB-B-DB (42 sites per mouth) protocols, respectively.

For the extent of AL, the relative bias ranged from -32.1% to 50.0% (Table 1). Among the fifteen PMPE protocols, the r6sites, fMB-DB-ML-DL, and rMB-DB-ML-DL protocols performed better, with a relative bias of <5.0% in absolute value for extent of AL ≥2, 4 and 6 mm. The rMB-B-DL, fMB-DL and rMB-DL protocols produced a relative bias of ≤5.0% in absolute value for AL ≥2 or 4 mm, while the relative bias increased to -17.9% to -7.1% in absolute value for extent of AL ≥6 mm. Other protocols with a relative bias <20.0% in absolute value included rMB-B-DB, fMB-B-DL, fMB-DB, rMB-DB and RT. The largest overestimation was observed in the CPI, where the relative bias for AL ≥6mm equaled 50.0% ($P < 0.001$). Large underestimations were observed for the fMB-B and rMB-B protocols, where the relative biases for the extent of AL

≥6 mm were -28.6% and -32.1% ($P < 0.001$), respectively.

For the extent of PD, the situation was similar to the extent of AL. Again, r6sites performed best. The relative bias of r6sites was 0 for extent of PD≥4mm or 6mm. However, interproximal sites protocols, such as fMB-DB, rMB-DB, fMB-DL, rMB-DL, fMB-DB-ML-DL, and rMB-DB-ML-DL caused severe overestimation, with the relative bias ranging from 12.5% to 54.2% (Table 2).

Regarding the extent of BOP (Table 2), the least biased result was, again, seen in r6sites, with a relative bias of only 0.3% ($P = 0.523$), followed by fMB-DB and rMB-DB (0.6% and -0.6%, respectively). The largest underestimation was for the fMB-B and rMB-B (-19.4% and -19.3%, respectively). CPI, fMB-DB-ML-DL, and rMB-DB-ML-DL overestimated the BOP extent by ≥10.0% ($P < 0.001$).

Bias for severity estimates

Table 3 shows sample severity estimates and relative biases for the various PMPE protocols. In a manner similar to the extent estimates, the relative biases were either negative or positive, indicating under- or overestimation. Protocols that examined different numbers of the same type of sites showed almost identical mean values. The r6sites, rMB-B-DB, fMB-B-DL, and rMB-B-DL protocols showed a relative bias of <5.0% in absolute value for both PD and AL severity.

For AL severity, relative biases ranged from -10.2% to 14.4% (Table 3). Among the fifteen PMPE protocols, r6sites performed best, with a relative bias of 0 ($P = 0.787$), followed by

fMB-DL and rMB-DL, with relative biases of 1.1% (P = 0.087 and 0.128, repectively) (Table 3). The fMB-B and rMB-B protocols underestimated the mean AL by 10.2% (P < 0.001). The RT and CPI protocols overestimated mean AL by 7.5% and 14.4%, respectively (P < 0.001).

Regarding relative biases for the severity of PD (Table 3), again, r6sites performed best, with a relative bias of 0 (P = 0.881). RT presented better performance in PD than in AL estimates, with a relative bias of only -0.7% (P = 0.078). Interproximal sites protocols and CPI caused >10.0% overestimation.

Agreement between PMPE and FMPE protocols

Table 4 shows the ICCs for extent and severity estimates, and most ICCs were larger than 0.800. As the severity thresholds increased, the ICCs decreased. In addition, when the same type of sites was examined in non-index-teeth protocols, full-mouth protocols generally presented higher ICCs than their random half-mouth versions. R6sites showed an ICC of >0.950 across all the parameters examined, except for an extent of PD ≥6 mm (0.936). FMB-B-DL also showed an ICC of >0.950 in all evaluated parameters with the exception of the extent of BOP (0.939), followed by rMB-B-DL. The fMB-B and rMB-B protocols produced lower ICCs across all parameters, and neither had an ICC >0.950. The CPI protocol also produced lower ICCs, ranging from 0.646 to 0.924. Interproximal sites protocols showed high ICCs in AL extent (0.899-0.987) and severity (0.957-0.984), while for PD, the ICCs were lower (0.687-0.920 for extent; 0.635-0.728 for severity).

Figure 1 (A-O) presents Bland-Altman plots for individual extent estimates for AL ≥4 mm. The individual differences between the gold-standard and each PMPE protocol were plotted

against the mean of these values. PMPE protocols that had better agreement with the FMPE method were indicated by difference points distributed closer to Y = 0 and fluctuation around the Y = 0 line in a smaller range. We believe that individual differences within 20.0% and -20.0% of the gold-standard results were clinically acceptable as was also depicted in the Bland-Altman plots. Non-index-teeth protocols that examined different numbers within the same set of sites showed a similar distribution pattern of points in the Bland-Altman plots, while the variation in random half-mouth protocols was slightly larger than in the full-mouth versions. Points for index-teeth protocols were mostly distributed above Y = 0, indicating overestimation, and the variations became larger as the extent level increased. Points for fMB-B and rMB-B were mostly distributed below Y = 0, indicating underestimation. In other protocols, most points distributed within ±20.0%, and distributed equally around Y =0.

For the extent of PD ≥4 mm (Fig 2 A-O), r6sites, MB-B-DL and MB-B-DB protocols showed better agreement. Most points for the fMB-B and rMB-B protocols were below Y = 0, indicating underestimation. RT showed good agreement for PD ≥4 mm extent, as opposed to AL overestimation. Most points in interproximal sites protocols and CPI were above Y=0, indicating overestimation.

Supplementary Fig1 (A-O) in the online *Journal of Periodontology* shows the Bland-Altman plots for BOP extent. Better agreement was observed for the r6sites and RT protocols. The r6sites protocol produced smaller variation than did RT. Most points from MB-B, MB-B-DB and MB-B-DL protocols were distributed below Y = 0, indicating underestimation. Moreover, the magnitude and variation of underestimation increased as the BOP extent decreased. For the CPI, MB-DL, and MB-DB-ML-DL protocols, most points were distributed above Y = 0, indicating

overestimation.

The Bland-Altman plots for severity of AL (see Supplementary Fig 2 A-O in the online *Journal of Periodontology*) showed that the r6sites, MB-B-DB, MB-B-DL, MB-DB, MB-DL, and MB-DB-ML-DL protocols were in good agreement, with most difference points within ± 10.0% of the gold-standard results. Most points in the MB-B protocols were below Y = 0, indicating underestimation. The CPI and RT protocols showed evident overestimation and larger variation with an increasing severity level. Bland-Altman plots for mean PD showed that the r6sites, MB-B-DB, MB-B-DL, and RT protocols showed good agreement, with nearly all individual differences between ± 10.0% of the gold-standard results (see Supplementary Fig 3 A-O in the online *Journal of Periodontology*). This was particularly true for the r6sites protocol. Most points for the MB-B protocols were below Y = 0, indicating underestimation. CPI and interproximal sites protocols caused severe overestimation.

Discussion

In the present study, we examined the accuracy of extent and severity estimates of fifteen PMPE protocols in a Chinese population with a range of periodontal diseases. Our findings showed that, of the PMPE protocols investigated in this study, the random half-mouth six-sites protocol produced the smallest bias and had the best agreement for AL, PD and BOP extent and severity estimates.

The relative bias for r6sites were all <5.0% in absolute value, for both severity and extent estimates. In addition, r6sites provided good agreement with FMPE data. Notably, for extent estimates, it provided excellent accuracy at AL/PD ≥4 mm and AL/PD ≥6 mm for assessments of

both bias and agreement. This indicated that the r6sites protocol should be suitable for the description of periodontitis extent in both mild and more advanced cases. Results with a low bias (<5.0% in absolute value) for the r6sites protocol and AL/PD ≥4 mm have also been reported in a rural Guatemalan population[24] and patients from Brazilian hospitals[29]. For AL/PD ≥6 mm, comparable results were even fewer, and a study by Dowsett et al. showed relative biases for the r6sites protocol of -6.7% and 2.8%, respectively[24].

Furthermore, r6sites also produced the most accurate results for BOP, an important parameter in periodontal inflammation that has been recommended for recording in epidemiological surveys[39]. To the best of our knowledge, the accuracy of BOP extent in PMPE protocols had not been investigated in previous studies. Gingival index, another indicator of gingival inflammation, had been assessed previously, and half-mouth six-sites per tooth protocol were also reported to show good reliability[24]. Combining limited bias and good agreement for the extent and severity of AL, PD and BOP, r6sites protocol is a suitable protocol for describing the extent and severity of periodontal disease in future epidemiologic studies.

Our results also showed that the fMB-B-DL and rMB-B-DL protocols performed very well, especially for the mean AL/PD and extent estimates of less severe cut-off values (AL/PD ≥4 mm). In the severity of AL and PD, these two protocols produced <5.0% relative bias in absolute value. Kingman et al., in a large general-population sample from Brazil, also demonstrated that these two protocols produced very small relative biases (<1.0% in absolute value for AL and -1.3% for PD) [5], but their study contained no information about the accuracy of extent estimates. In our data regarding extent estimates of AL ≥2 and 4 mm, PD ≥4 mm and BOP, the relative biases were generally <5.0% in absolute value. However, in more severe cases such as AL/PD ≥6 mm, the

relative biases increased to <15.0% in absolute value. These results suggested that MB-B-DL protocols are more suitable for the description of severity and extent in mild rather than advanced cases.

For other non-index-teeth protocols, our findings showed that full- and random half-mouth protocols examining different numbers of the same sites produced similar levels of relative bias. For example, the fMB-B-DB and rMB-B-DB protocols produced close to the same mean score and relative bias. However, differences in relative bias were greater for protocols examining different types of sites. For example, the r6sites, fMB-B-DB, rMB-B-DB, fMB-B-DL and rMB-B-DL protocols generally showed <20.0% relative bias for extent and <5.0% relative bias for severity (in absolute value). We would like to note that the ratio of interproximal to mid-facial sites was 2:1 in all of these low biased protocols, which was the same as the gold-standard protocol. However, extent relative bias was over 20.0% and severity relative bias was over 10.0% for the MB-B protocols (interproximal to mid-facial sites 1:1). For interproximal sites protocols, such as MB-DB, MB-DL, and MB-DB-ML-DL, they showed good accuracy in AL extent and severity, but they overestimated extent and severity of PD and BOP. This indicated that bias in extent and severity estimates may be more directly related to the type of sites examined, rather than the number. Particularly, a ratio of interproximal to mid-facial sites of 2:1 in non-index-teeth protocols (the same as FMPE) may be essential for the accuracy of extent and severity estimates from PMPE protocols.

In index-teeth protocols, the use of CPI teeth caused considerable overestimation and provided the poorest extent and severity estimates. The tooth types selected in the CPI protocol may explain the large overestimation, considering that molars and incisors are more susceptible to

periodontal disease, and furcation involvement in molars increased AL, PD, and BOP[20]. In our results for AL and PD severity, overestimation was larger than 10.0%. A study of a Brazilian population also showed an overestimation of CPI in which the relative biases for mean PD and AL were 12.7% and 14.8%, respectively[29]. For extent of AL/PD ≥4 mm, overestimation was 28.6% and 49.6%, respectively, in our results. Vettore et al. and Relvas et al. also found that CPI overestimated the extent scores[29, 35]. For PD ≥4 mm and AL ≥4mm, overestimation in these studies were both larger than 35.0%[29, 35]. For AL/PD ≥6 mm, CPI caused an even greater overestimation of 50.0% and 108.3%, respectively, in our data. Vettore et al. reported a similar overestimation (>60.0%) at AL/PD >6 mm[29]. These results indicated that CPI might not be suitable for the determination of extent and severity estimates.

RT, which has been assessed frequently in previous studies[2, 12, 24, 26, 29], provided small biases and good agreements in severity and extent of PD and BOP in our study, but this method performed not as well as the r6sites protocol. Furthermore, RT produced a relatively large overestimation of the extent and severity of AL (18.6% for AL ≥4 mm extent and 7.5% for severity). Since AL is the sum of PD and recession, we hypothesized that the overestimation of recession in RT contributed to the large bias for AL. To prove this, we calculated the relative bias of AL for recession, and found that it overestimated the extent of recession ≥3 mm and severity by 17.2% and 17.0%, respectively. The overestimation of AL in RT has been reported in a rural Guatemalan population (1.6% for AL ≥4 mm, and 13.9% for AL ≥6 mm)[24] and a Portuguese population (13.9% for AL ≥4 mm, and 18.6% for mean AL)[35], which supported our results.

Limitations

The results of this study were derived from a population with a range of different periodontal diseases. The relevance of these results to other populations will require further confirmation. In addition, the sample size was small. Studies with larger sample sizes will be needed to further evaluate the accuracy of extent and severity estimates for the various PMPE protocols.

In the present study, we focused only on the accuracy of extent and severity estimates of PMPE protocols. That PMPE protocols cause biases in prevalence estimates[34, 35] has been well documented. To address this issue, some researchers have proposed inflation factors to adjust for prevalence bias[27, 30, 31]. An ongoing pilot study is being conducted to analyze the bias of prevalence estimates and accuracy of adjusted prevalence estimates with inflation factors.

Conclusion

Within the limitation of this study, the r6sites protocol produced very accurate extent and severity estimates for AL, PD, and BOP when compared with the FMPE protocol. When FMPE is not feasible, we recommend using the r6sites protocol when a comprehensive description of disease extent and severity is required.

Acknowledgements

The cross-sectional epidemiological survey performed in 2000 in Beijing for periodontitis was supported by Procter & Gamble Corporation, Cincinnati, Ohio, USA.

Conflict of interest statement

The authors report no conflicts of interest with this study.

Ethics Statement

The dataset of the study "Accuracy of Partial-Mouth Examination Protocols for Extent and Severity Estimates of Periodontitis: A Study in a Chinese Population with Chronic Periodontitis" came from the data of a cross-sectional epidemiological survey performed in Beijing, in the year 2000. This survey performed in 2000 was approved by the institutional review board of Procter & Gamble Corporation, and written informed consent was obtained from each subject in accordance with the Declaration of Helsinki. We re-analyzed the data of this 2000 survey for accuracy of partial-mouth examination protocols, and this re-analysis study was approved by the institutional review board of Peking University School and Hospital of Stomatology (PKUSSIRB-201413037).

Footnotes

‡ UNC-15, Hu-Friedy, Chicago, USA

§ SPSS Version 16.0, Chicago IL, USA

‖ GraphPad Prism version 5.01 for Windows, San Diego CA, USA

References

1. Beck JD, Loe H. Epidemiological principles in studying periodontal diseases. *Periodontol 2000* 1993;2:34-45.
2. Kingman A, Albandar JM. Methodological aspects of epidemiological studies of periodontal diseases. *Periodontol 2000* 2002;29:11-30.
3. Owens JD, Dowsett SA, Eckert GJ, Zero DT, Kowolik MJ. Partial-mouth assessment of periodontal disease in an adult population of the United States. *J Periodontol* 2003;74:1206-1213.
4. Benigeri M, Brodeur JM, Payette M, Charbonneau A, Ismail AI. Community periodontal index of treatment needs and prevalence of periodontal conditions. *J Clin Periodontol* 2000;27:308-312.
5. Kingman A, Susin C, Albandar JM. Effect of partial recording protocols on severity estimates of periodontal disease. *J Clin Periodontol* 2008;35:659-667.
6. Ramfjord SP. Indices of prevalence and incidence of periodontal disease. *J Periodontol* 1959;30:51-59.
7. Ainamo J, Barmes D, Beagre G, Cutress T, Martin J, Sardo-Infirri J. Development of the World Health Organization (WHO) community periodontal index of treatment needs (CPITN). *Int Dent J* 1982;32:281-291.
8. Hunt RJ, Levy SM, Beck JD. The prevalence of periodontal attachment loss in an Iowa population aged 70 and older. *J Public Health Dent* 1990;50:251-256.
9. Papapanou PN, Wennstrom JL, Johnsson T. Extent and severity of periodontal destruction based on partial clinical assessments. *Community Dent Oral Epidemiol* 1993;21:181-184.
10. Albandar JM, Brunelle JA, Kingman A. Destructive periodontal disease in adults 30 years of age and older in the United States, 1988-1994. *J Periodontol* 1999;70:13-29.
11. Thomson WM, Hashim R, Pack AR. The prevalence and intraoral distribution of periodontal attachment loss in a birth cohort of 26-year-olds. *J Periodontol* 2000;71:1840-1845.
12. Beck JD, Caplan DJ, Preisser JS, Moss K. Reducing the bias of probing depth and attachment level estimates using random partial-mouth recording. *Community Dent Oral Epidemiol* 2006;34:1-10.
13. Brennan DS, Spencer AJ, Roberts-Thomson KF. Periodontal disease among 45-54 year olds in Adelaide, South Australia. *Aust Dent J* 2007;52:55-60.
14. Tran DT, Gay I, Du XL, et al. Assessment of partial-mouth periodontal examination protocols for periodontitis surveillance. *J Clin Periodontol* 2014;41:846-852.
15. Dye BA, Thornton-Evans G. A brief history of national surveillance efforts for periodontal disease in the United States. *J Periodontol* 2007;78:1373-1379.
16. Ainamo J, Ainamo A. Partial indices as indicators of the severity and prevalence of periodontal disease. *Int Dent J* 1985;35:322-326.
17. Hunt RJ. The efficiency of half-mouth examinations in estimating the prevalence of periodontal disease. *J Dent Res* 1987;66:1044-1048.
18. Kingman A, Morrison E, Loe H, Smith J. Systematic errors in estimating prevalence and severity of periodontal disease. *J Periodontol* 1988;59:707-713.
19. Hunt RJ, Fann SJ. Effect of examining half the teeth in a partial periodontal recording of older adults. *J Dent Res* 1991;70:1380-1385.
20. Baelum V, Fejerskov O, Manji F, Wanzala P. Influence of CPITN partial recordings on estimates of prevalence and severity of various periodontal conditions in adults. *Community Dent Oral Epidemiol* 1993;21:354-359.
21. Agerholm DM, Ashley FP. Clinical assessment of periodontitis in young adults--evaluation of probing depth and partial recording methods. *Community Dent Oral Epidemiol* 1996;24:56-61.
22. Eaton KA, Duffy S, Griffiths GS, Gilthorpe MS, Johnson NW. The influence of partial and full-mouth recordings on estimates of prevalence and extent of lifetime cumulative attachment loss: a study in a population of young male military recruits. *J Periodontol* 2001;72:140-145.

23. Thomson WM, Williams SM. Partial- or full-mouth approaches to assessing the prevalence of and risk factors for periodontal disease in young adults. *J Periodontol* 2002;73:1010-1014.
24. Dowsett SA, Eckert GJ, Kowolik MJ. The applicability of half-mouth examination to periodontal disease assessment in untreated adult populations. *J Periodontol* 2002;73:975-981.
25. Borges-Yanez SA, Maupome G, Jimenez-Garcia G. Validity and reliability of partial examination to assess severe periodontitis. *J Clin Periodontol* 2004;31:112-118.
26. Mumghamba EG, Pitiphat W, Matee MI, Simon E, Merchant AT. The usefulness of using Ramfjord teeth in predicting periodontal status of a Tanzanian adult population. *J Clin Periodontol* 2004;31:16-18.
27. Susin C, Kingman A, Albandar JM. Effect of partial recording protocols on estimates of prevalence of periodontal disease. *J Periodontol* 2005;76:262-267.
28. Bassani DG, da Silva CM, Oppermann RV. Validity of the "Community Periodontal Index of Treatment Needs" (CPITN) for population periodontitis screening. *Cad Saude Publica* 2006;22:277-283.
29. Vettore MV, Lamarca Gde A, Leao AT, Sheiham A, Leal Mdo C. Partial recording protocols for periodontal disease assessment in epidemiological surveys. *Cad Saude Publica* 2007;23:33-42.
30. Eke PI, Thornton-Evans GO, Wei L, Borgnakke WS, Dye BA. Accuracy of NHANES periodontal examination protocols. *J Dent Res* 2010;89:1208-1213.
31. Albandar JM. Underestimation of periodontitis in NHANES surveys. *J Periodontol* 2011;82:337-341.
32. Peres MA, Peres KG, Cascaes AM, et al. Validity of partial protocols to assess the prevalence of periodontal outcomes and associated sociodemographic and behavior factors in adolescents and young adults. *J Periodontol* 2012;83:369-378.
33. Relvas M, Diz P, Velazco C, Otero JL, Pacheco JJ, Tomas I. Evaluation of partial-mouth recording systems of gingival parameters in a Portuguese adult population. *J Public Health Dent* 2013;73:135-146.
34. Tran DT, Gay I, Du XL, et al. Assessing periodontitis in populations: a systematic review of the validity of partial-mouth examination protocols. *J Clin Periodontol* 2013;40:1064-1071.
35. Relvas M, Tomas I, Salazar F, Velazco C, Blanco J, Diz P. Reliability of partial-mouth recording systems to determine periodontal status: a pilot study in an adult Portuguese population. *J Periodontol* 2014;85:e188-197.
36. Ouyang Xiangying, Cao Caifang, Hu Wenjie, Wang Hongying, J Leslie Winston, Liu H. Survey on periodontal disease status in 200 chronic periodontitis patients (in Chinese). *J Modern Stomatol* 2004;4:355-358.
37. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods* 1996;1:30-46.
38. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-310.
39. Leroy R, Eaton KA, Savage A. Methodological issues in epidemiological studies of periodontitis - how can it be improved? *BMC Oral Health* 2010;10:8-14.

Figure legends

Figure 1 Bland–Altman plots for extent estimates of AL ≥4 mm (N = 200). Bland–Altman plots (A-O) show the agreement pattern across disease level between each PMPE protocol and the gold-standard protocol. Black points represent differences in individual AL ≥4 mm extent estimates between each PMPE protocol and the gold-standard protocol. Blue and purple points represent 20.0% and -20.0% of the gold-standard results, respectively.

Figure 2 Bland–Altman plots for the extent estimates of PD ≥4 mm (N = 200). Bland–Altman plots (A-O) show the agreement pattern across disease level between each PMPE protocol and the gold-standard protocol. Black points represent differences in individual PD ≥4 mm extent estimates between each PMPE protocol and the gold-standard protocol. Blue and purple points represent 20.0% and -20.0% of the gold-standard results, respectively.

Table 1. Extent Estimates and Relative Biases of Attachment Loss for Partial-mouth Examination Protocols*

Protocols	AL≥2mm				AL≥4mm				AL≥6mm			
	Mean	SD	Relative bias (%)	*P* value	Mean	SD	Relative bias (%)	*P* value†	Mean	SD	Relative bias (%)	*P* value†
FMPE	51.6	24.9	-	-	14.0	17.3	-	-	2.8	5.6	-	-
Non-index-teeth protocols												
r6sites	51.4	25.1	**-0.4**	0.262†	14.0	17.4	**0**	0.843	2.9	5.7	**3.6**	0.578
fMB-B-DB	49.4	24.0	**-4.3**	<0.001	12.6	16.2	-10.0	<0.001	2.2	5.0	-21.4	<0.001
rMB-B-DB	49.9	24.1	**-3.3**	0.001	12.8	16.9	-8.6	<0.001	2.4	5.6	-14.3	0.007
fMB-B-DL	50.3	24.7	**-2.5**	<0.001	13.2	17.2	-5.7	<0.001	2.4	5.2	-14.3	<0.001
rMB-B-DL	50.8	25.3	**-1.6**	0.064	13.5	17.8	**-3.6**	0.013	2.6	5.7	-7.1	0.099
fMB-B	47.0	24.6	-8.9	<0.001†	11.8	16.3	-15.7	<0.001	2.0	5.0	-28.6	<0.001
rMB-B	46.7	25.3	-9.5	<0.001	11.8	16.8	-15.7	<0.001	1.9	5.2	-32.1	<0.001
fMB-DB	51.2	24.9	**-0.8**	0.415	12.5	16.3	-10.7	<0.001	2.3	5.0	-17.9	0.001
rMB-DB	51.1	25.7	**-1.0**	0.494	12.2	16.4	-12.9	<0.001	2.5	5.7	-10.7	0.010
fMB-DL	52.5	26.0	**1.7**	0.016	13.4	18.0	**-4.3**	0.003	2.5	5.5	-10.7	0.008
rMB-DL	52.7	26.4	**2.1**	0.029	13.3	18.2	**-5.0**	0.011	2.3	5.4	-17.9	0.003
fMB-DB-ML-DL	53.3	25.6	**3.3**	<0.001	13.6	17.7	**-2.9**	0.126	2.7	5.6	**-3.6**	0.148
rMB-DB-ML-DL	53.2	26.1	**3.1**	0.001	13.6	18.0	**-2.9**	0.129	2.7	5.5	**-3.6**	0.555
Index-teeth protocols												
RT	55.3	25.3	7.2	<0.001	16.6	19.5	18.6	<0.001	3.1	6.4	10.7	0.313
CPI	59.0	25.0	14.3	<0.001	18.0	20.9	28.6	<0.001	4.2	7.8	50.0	<0.001

Table 2. Extent Estimates and Relative Biases of Probing Depth for Partial-mouth Examination Protocols*

Protocols	PD≥4mm				PD≥6mm				BOP			
	Mean	SD	Relative bias (%)	*P* value	Mean	SD	Relative bias (%)	*P* value†	Mean	SD	Relative bias (%)	*P* value
FMPE	25.2	13.9	-	-	2.4	3.7	-	-	71.0	15.7	-	-
Non-index-teeth protocols												
6sites	25.2	14.3	**0**	0.841	2.4	3.9	**0**	0.748	71.2	16.1	**0.3**	0.523
fMB-B-DB	24.7	14.4	**-2.0**	0.145	2.5	4.1	**4.2**	0.061	62.6	18.8	-11.8	<0.001
rMB-B-DB	24.8	14.8	**-1.6**	0.402	2.4	4.2	**0**	0.754	62.7	19.8	-11.7	<0.001
fMB-B-DL	24.2	13.8	**-4.0**	<0.001†	2.1	3.8	-12.5	<0.001	66.8	16.4	-5.9	<0.001
rMB-B-DL	24.3	14.0	**-3.6**	0.001	2.0	4.1	-16.7	0.013	67.1	17.0	-5.5	<0.001
fMB-B	19.5	12.6	-22.6	<0.001	1.7	3.3	-29.2	<0.001	57.2	19.4	-19.4	<0.001†
rMB-B	19.5	13.1	-22.6	<0.001	1.7	3.6	-29.2	<0.001	57.3	20.6	-19.3	<0.001
fMB-DB	34.7	19.9	37.7	<0.001†	3.5	5.8	45.8	<0.001	71.4	19.2	**0.6**	0.575
rMB-DB	35.3	21.5	40.1	<0.001	3.7	6.5	54.2	<0.001	70.6	21.3	**-0.6**	<0.001
fMB-DL	33.9	18.6	34.5	<0.001†	2.9	5.3	20.8	<0.001	77.7	15.6	9.4	<0.001
rMB-DL	33.8	19.3	34.1	<0.001†	2.7	5.3	12.5	0.821	77.3	16.4	8.9	<0.001
fMB-DB-ML-DL	34.9	18.1	38.5	<0.001†	3.2	5.2	33.3	<0.001	78.4	15.3	10.4	<0.001
rMB-DB-ML-DL	35.0	18.3	38.9	<0.001	3.1	5.0	29.2	<0.001	78.1	16.2	10.0	<0.001
Index-teeth protocols												
RT	25.3	14.7	**0.4**	0.839	2.2	4.4	-8.3	0.135	69.6	16.5	**-2.0**	0.001
CPI	37.7	15.6	49.6	<0.001	5.0	6.0	108.3	<0.001	78.3	14.2	10.3	<0.001

*Boldface numbers have <5.0% relative bias in absolute value.

†Differences between PMPE and FMPE protocols were tested with the Wilcoxon-signed rank test, and others were tested with the paired t-test.

Table 3. Mean Values and Relative Biases of Attachment Loss and Probing Depth for Partial-mouth Examination Protocols*

Protocols	AL				PD			
	Mean	SD	Relative bias (%)	*P* value	Mean	SD	Relative bias (%)	*P* value†
FMPE	1.87	0.93	-	-	2.93	0.41	-	-
Non-index protocols								
r6sites	1.87	0.94	**0**	0.787†	2.93	0.42	**0**	0.881
fMB-B-DB	1.77	0.89	-5.3	<0.001†	2.88	0.43	**-1.7**	<0.001
rMB-B-DB	1.79	0.90	**-4.3**	<0.001	2.89	0.44	**-1.4**	<0.001
fMB-B-DL	1.81	0.91	**-3.2**	<0.001	2.86	0.40	**-2.4**	<0.001†
rMB-B-DL	1.83	0.94	**-2.1**	0.005	2.86	0.42	**-2.4**	<0.001
fMB-B	1.68	0.89	-10.2	<0.001†	2.66	0.40	-9.2	<0.001†
rMB-B	1.68	0.91	-10.2	<0.001†	2.66	0.42	-9.2	<0.001
fMB-DB	1.83	0.89	**-2.1**	0.051†	3.31	0.52	13.0	<0.001
rMB-DB	1.82	0.89	**-2.7**	0.057†	3.32	0.55	13.3	<0.001
fMB-DL	1.89	0.94	**1.1**	0.087	3.28	0.47	11.9	<0.001
rMB-DL	1.89	0.94	**1.1**	0.128	3.27	0.48	11.6	<0.001
fMB-DB-ML-DL	1.92	0.93	**2.7**	<0.001	3.31	0.47	13.0	<0.001
rMB-DB-ML-DL	1.91	0.95	**2.1**	0.002	3.30	0.48	12.6	<0.001
Index-teeth protocols								
RT	2.01	0.97	7.5	<0.001	2.91	0.44	**-0.7**	0.078
CPI	2.14	1.04	14.4	<0.001†	3.33	0.46	13.7	<0.001†

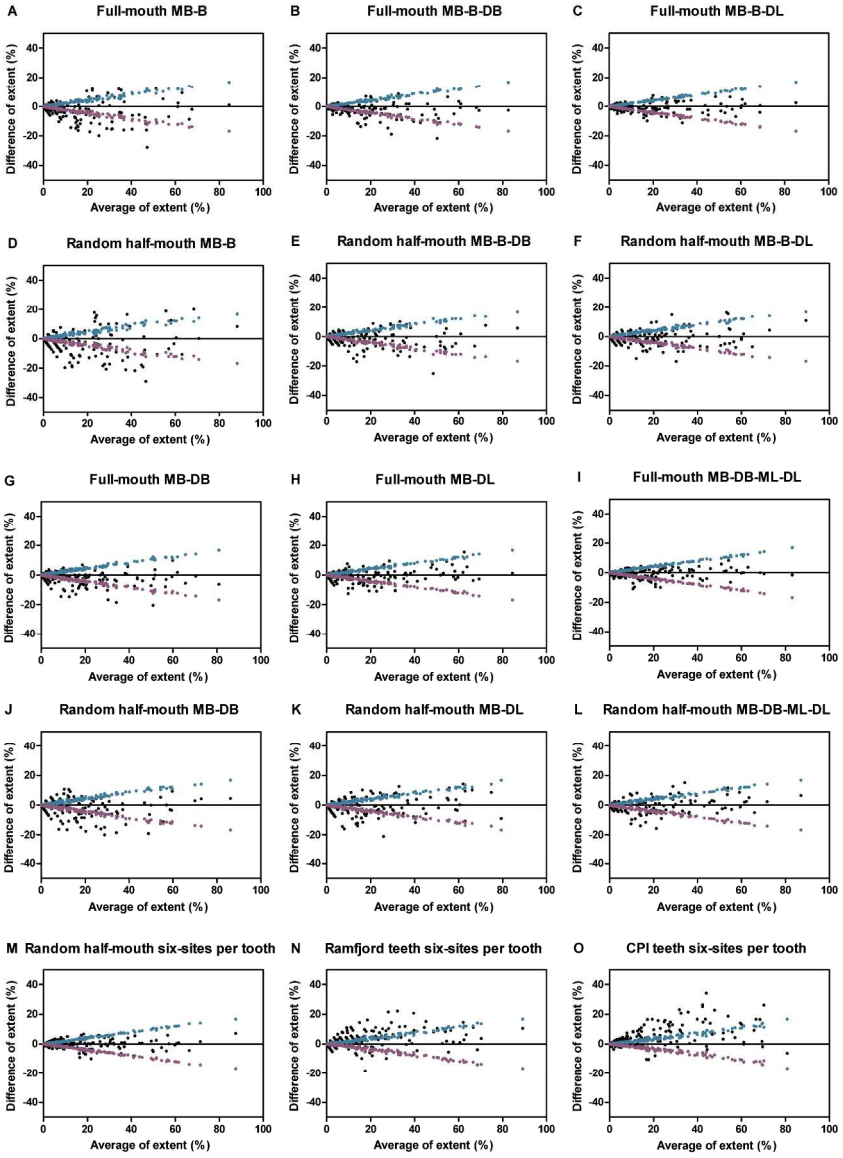
*Boldface numbers have <5.0% relative bias in absolute value.

†Differences between PMPE and FMPE protocols were tested with the Wilcoxon-signed rank test, and others were tested with the paired t-test.

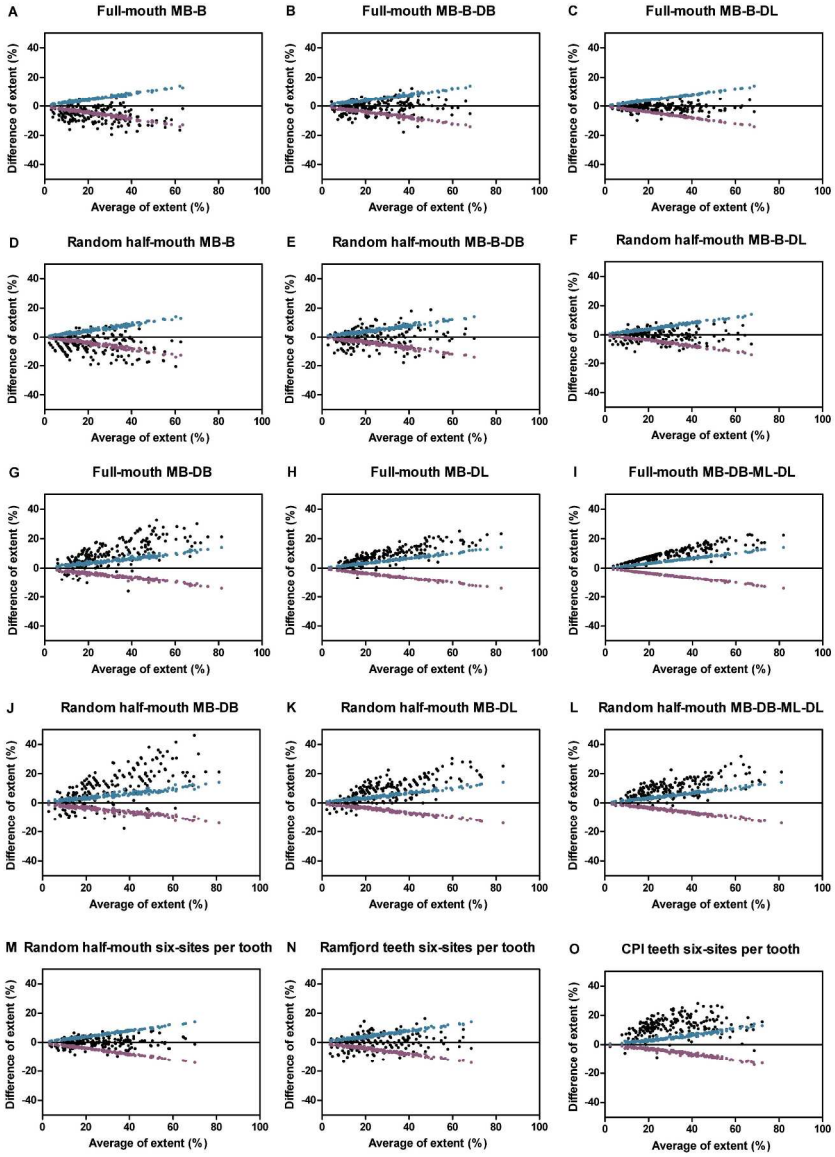
Table 4. Intra-class Correlation Coefficients for Partial-Mouth Examination Protocols*

Protocols	Extent						Severity	
	AL≥2mm	AL≥4mm	AL≥6mm	PD≥4mm	PD≥6mm	BOP	AL	PD
Non-index- teeth protocols								
r6sites	**0.989**	**0.988**	**0.962**	**0.975**	0.936	**0.965**	**0.992**	**0.982**
fMB-B-DB	**0.971**	**0.966**	0.919	0.948	0.947	0.826	**0.972**	**0.953**
rMB-B-DB	**0.956**	**0.958**	0.874	0.908	0.872	0.788	**0.964**	0.927
fMB-B-DL	**0.989**	**0.986**	**0.968**	**0.982**	**0.965**	0.939	**0.989**	**0.975**
rMB-B-DL	**0.974**	**0.968**	0.935	**0.954**	0.892	0.900	**0.979**	**0.952**
fMB-B	0.939	0.931	0.833	0.853	0.896	0.686	**0.938**	0.772
rMB-B	0.919	0.899	0.838	0.822	0.815	0.663	**0.924**	0.753
fMB-DB	**0.957**	**0.963**	0.911	0.752	0.830	0.892	**0.972**	0.676
rMB-DB	0.916	0.941	0.899	0.687	0.726	0.824	**0.957**	0.635
fMB-DL	**0.977**	**0.979**	**0.961**	0.822	0.891	0.881	**0.984**	0.728
rMB-DL	**0.958**	**0.959**	0.915	0.796	0.810	0.851	**0.974**	0.719
fMB-DB-ML-DL	**0.982**	**0.987**	**0.975**	0.810	0.920	0.872	**0.987**	0.707
rMB-DB-ML-DL	**0.967**	**0.972**	0.934	0.783	0.874	0.854	**0.978**	0.700
Index-teeth protocols								
RT	**0.956**	0.944	0.897	0.932	0.857	0.925	**0.964**	0.948
CPI	0.918	0.917	0.852	0.667	0.720	0.814	0.924	0.646

*Boldface numbers have >0.950 intra-class correlation coefficients.



231x309mm (300 x 300 DPI)



234x314mm (300 x 300 DPI)