

ABSTRACT

Examiner inexperience, particularly with automated periodontal probes, may contribute to increased variability in probing depth measurement. To assess the impact of experience on probing, we compared Florida probe pocket depth (PD) measurements made by a standardized and highly experienced examiner to those made by two naive examiners who received limited training. Study objectives were to assess 1) intra-examiner variability of replicate pocket depth measurements and 2) inter-examiner consistency of pocket depths. Eligibility in this randomized, split-mouth study was limited to 12 adult volunteers who had at least one periodontal pocket of 5-10 mm on each side of the mouth. For each subject, the three teeth with the deepest pockets on each side of the mouth were selected for probing. The standard examiner made duplicate PD measurements at six sites on all six test teeth in the mouth, while test examiners made similar measurements on the three test teeth in a randomly assigned half mouth. Examination order was randomized, and all examiners were blinded as to measurement outcomes. Results from this probing study showed the measurement variability estimates for the two test examiners to be 0.374 and 0.226 with 90% confidence intervals of (0.321, 0.441) and (0.194, 0.267), respectively. Corresponding estimates and intervals for the standard examiner were 0.370, 0.418, (0.318, 0.437) and (0.359, 0.495), respectively. One-sided p-values for comparing variability of test examiners to the standard were 0.954 and 0.996. The mean PD differences between the test and standard examiners were 0.02 mm ($p=0.799$) and 0.02 mm ($p=0.747$). **Both test examiners were found to be repeatable, having mean pocket depth measurements similar to and variability not exceeding that of the standard examiner.**

INTRODUCTION

Current probing methods of measuring periodontal pocket depths at specific sites are reported to have substantial intra-examiner and inter-examiner variability, which decrease measurement precision.¹ Excessive measurement variability can obscure site-specific changes.² Primary sources of such variability include: examiner variability, subject/site variability, and instrument variability.³

Examiner training and calibration may reduce examiner variability attributed to probe misangulation and malposition. Both should be conducted prior to clinical trials involving probing measurements.⁴ First, the level of intra-examiner measurement error for individual examiners must be established in order to set diagnostic threshold values for monitoring change.⁴ Second, examiner training and calibration may result in decreased measurement variability which increases the precision of the measurement and the statistical power of the study.^{1,3}

OBJECTIVES

To assess the intra-examiner variability for the measurement of periodontal pocket depths via replicate measurements. To assess the inter-examiner consistency of measured periodontal pocket depths between a standard examiner and two test examiners.

MATERIALS AND METHODS

Study Design

This study utilized a randomized, single center design on 12 adults with periodontal disease, including at least two periodontal pockets with probing depths of 5-10 mm, one on each side of the mouth. For each subject, the three teeth with the deepest manual probing pocket depths on each side of the mouth were selected, resulting in six test teeth. A standard examiner and two test examiners examined each subject. The standard examiner measured all six test teeth and the test examiners each measured three teeth, depending on which side of the mouth they were randomized to measure for each particular subject (Table 1). Replicate measurements of pocket depth were taken with a controlled pressure periodontal probe.

Table 1. Measurement Assignment

First Examination	
Right Side	Left Side
Standard Examiner Standard Examiner Test Examiner 1 Test Examiner 2	Standard Examiner Standard Examiner Test Examiner 2 Test Examiner 1
Second Examination	
Right Side	Left Side
Test Examiner 1 Test Examiner 2 Standard Examiner Standard Examiner	Test Examiner 2 Test Examiner 1 Standard Examiner Standard Examiner

Statistical Models

Objective 1. Individual Examiner Measurement Error Variances

$Y_{ijk} = \mu + a_i + \beta_{j(i)} + e_{ij}$
<p>Where: $i = 1, \dots, 12$ is an index for patients $j = 1, \dots, 18$ is an index for sites within patient k is an index for rep Y_{ijk} is the response</p> <p>Fixed Effects: μ is the grand mean</p> <p>Mutually Independent Normal Random Variables: a_i represents subject variability $\beta_{j(i)}$ represents site within subject variability e_{ij} represents measurement error These random variables have zero means and variances s_a^2, s_b^2, and s_e^2, respectively</p>

A 90% confidence interval for pocket depth measurement error (s^2) for each test examiner was calculated via standard techniques for variance estimation⁵. The upper limit of this confidence interval was compared to 0.55 mm². If this limit was less than 0.55 mm² then the test examiner was to be considered qualified.

To compare examiners with respect to mean pocket depths, the statistical model above was used where the response represents pocket depth measurement. The test for examiner differences in mean pocket depth score was a two-sided test with a 5% type 1 error rate. Patient by examiner and site by examiner variance components were estimated via the method of moments.

RESULTS

Objective 2. Comparisons of Test Examiner to the Standard Examiner

$Y_{ijkl} = \mu + a_i + \beta_{j(i)} + E_K + a_i E_K + \beta_{j(i)} E_K + e_{ijk}$ <p>Where: $i = 1, \dots, 12$ is an index for patients $j = 1, \dots, 18$ is an index for sites within patient $k = 1, 2$ is an index for examiner $l =$ is an index for rep Y_{ijkl} is the response</p> <p>Fixed Effects: μ is the grand mean E_K is the examiner effect</p> <p>Mutually Independent Normal Random Variables: a_i represents subject variability $\beta_{j(i)}$ represents site within subject variability $a_i E_K$ represents subject by examiner interaction $\beta_{j(i)} E_K$ represents site by examiner interaction e_{ijk} represents measurement error These random variables have zero means and variances $s^2_a, s^2_\beta, s^2_{aE}, s^2_{\beta E}$ and s^2_e, respectively</p>
--

To compare examiners with respect to measurement error, the statistical model above was used where the response was the natural log of the variance of repeated measurements + small constant (0.0002). Repeatability of test examiners relative to the standard examiner with respect to pocket depth measurements was assessed via a mixed model analysis of variance on the natural log⁶ of the variability of the repeated measurements at each site. The test for examiner differences was one-sided with the alternative being greater measurement variability for the test examiner than for the standard examiner.

Objective 1:

The 90% confidence intervals for measurement error for both test examiners were entirely below 0.55 mm² (Table 2, Figures 1&2). This was also true for the standard examiner for both sets of periodontal sites examined.

Table 2. Variance of Replicate Pocket Depth Measurements

Sites Common To	Examiner	# Sites	Estimate	90% Confidence Level	
				Lower Limit	Upper Limit
Standard Examiner & Test Examiner 1	Standard Examiner & Test Examiner 1	215	0.370	0.318	0.437
		215	0.374	0.321	0.441
Standard Examiner & Test Examiner 2	Standard Examiner & Test Examiner 2	212	0.418	0.359	0.495
		212	0.226	0.194	0.267

Figure 1: Initial and Replicate Pocket Depth Measurements

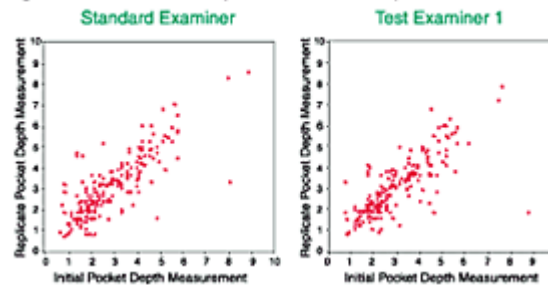
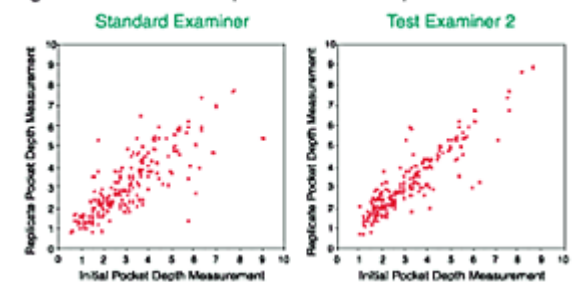


Figure 2: Initial and Replicate Pocket Depth Measurements



Objective 2:

Neither test examiners' variability was shown to be statistically significantly greater than that of the standard examiner for corresponding sites ($p=0.954$ and 0.996). The mean pocket depths for test examiner 1 and the standard examiner were 2.96 and 2.94 ($p=0.799$). The mean pocket depths for test examiner 2 and the standard examiner were 3.05 and 3.03 ($p=0.747$). For sites common to the Standard Examiner and Test Examiner 1, the patient by examiner and site by examiner variance components were estimated to be 0.017 and 0.223, respectively. For the Standard Examiner and Test Examiner 2, the patient by examiner and site by examiner variance components were estimated to be 0.001 and 0.311, respectively.

CONCLUSION

-Both test examiners were found to be sufficiently repeatable with respect to duplicate probing measurements according to the criteria established prior to the exercise.

-The probing variabilities of the two test examiners were not significantly greater than that of the standard examiner and the mean pocket depth measurements of the test examiners were not significantly different from that of the standard examiner.

REFERENCES

1. Walsh TF, Saxby MS. Inter- and intra examiner variability using a standard and constant force periodontal probes. *J Clin Periodontol* 1989; 16: 140-143
2. Cohen ME. Detection of periodontal probing change by analysis distribution mean and skew. *J Periodont Res* 1995; 30: 81-87
3. Hulley SB, Cummings SR. Planning the measurements: Precision and accuracy. In "*Designing Clinical Research*." Baltimore: Williams and Wilkins, 1988: 31-41.
4. Pihlstrom BL. Measurement of attachment level in clinical trials: Probing methods. *J Periodontol* 1992; 63: 1072-1077.
5. Burdick RK, Graybill FA. *Confidence intervals on variance components*. New York: Marcel Decker, 1992: 80.
6. Scheffé, H. *The analysis of variance*, New York: Wiley, 1959.