Research

Inter–examiner Reliability of Salivary Diagnostic Tests in a Practice–based Research Network

Marilynn Rothen, RDH, BS; Joana Cunha–Cruz, DDS, PhD; Lloyd Mancl, PhD; Brian Leroux, PhD; Brooke Latzke Davis, MPH; Justin Coyne; Jane Gillette, DDS; Joel Berg, DDS, MS

Introduction

Patients' salivary characteristics, such as flow rate, pH and buffering capacity, may provide information about future caries risk.[1-3] Presently, there are several commercially available test kits and methods for measuring salivary characteristics. The salivary diagnostic tests available chair–side today evolved from a laboratory method developed by Ericsson for measuring the buffering capacity of saliva electrometrically.[4] Ericsson's test was correlated with an increase in dental caries and is still considered the standard today. The first chair–side method of assessing buffering capacity was developed and commercialized a couple of decades later. In this method, the laboratory pH meter was substituted with a liquid colorimetric pH indicator in a test tube to which a known volume of stimulated saliva was added.[5] This technique was further simplified for chair–side use to a pH indicator strip that was impregnated with acid.[6] To use the buffer strip, the clinician places a drop of the patient's saliva on the acid test pad and determines the color reading produced. Both methods only determine crude estimates of buffering capacity within the range of low (pH≤4.0), intermediate (pH=4.5 to 5.5) and high (pH≥6.0). Using this range, they have been compared to Ericsson's test and have been shown to be valid measurements of salivary buffering capacity in the low to intermediate range, with a tendency to underestimate the values in the high range (not considered clinically sig-

Abstract

Purpose: Patient salivary characteristics are being measured as part of the Northwest PRECEDENT (Practice–based REsearch Collaborative in Evidence–based DENTistry) study on caries risk assessment. Prior to the implementation of these salivary diagnostic tests in a practice–based cohort study, inter–examiner reliability was assessed for resting salivary pH, stimulated salivary flow rate, pH and buffering capacity.

Methods: An initial evaluation of inter–examiner reliability of the 4 salivary tests was conducted among 6 dental auxiliary examiners using a convenience sample of 40 dental students. An incomplete block design was used to assign samples to examiners (3 examiners per sample and 20 samples per examiner). Inter–examiner reliability testing was conducted on a patient population representative of the practice–based network in 4 member practices. Two dental assistants per practice independently conducted the salivary tests on samples provided by a random selection of 20 to 25 patients. A separate analysis was performed for each study. For each test, an inter–examiner reliability index was computed.

Results: Results from two studies are reported. In the first, stimulated salivary flow rate demonstrated excellent inter–examiner reliability, and resting salivary pH showed high inter–examiner reliability, while buffering capacity and stimulated salivary pH had moderate and very low inter–examiner reliability, respectively. In the second, inter–examiner reliability was excellent for the stimulated salivary flow rate and the resting salivary pH. The inter–examiner reliability for the stimulated salivary pH was also high and the stimulated salivary buffering capacity test had moderate reliability.

Conclusion: The small variance in stimulated salivary pH and buffering capacity in dental students may have artificially made the reliability appear low in the first attempt at inter–examiner reliability testing. In the second study, all 4 tests had an acceptable performance.

Keywords: Inter–examiner reliability, salivary diagnostic tests, dental practice–based research networks, caries risk assessment

This study supports the NDHRA priority area, **Health Promotion/Disease Prevention:** Validate and test assessment instruments/strategies/mechanisms that increase health promotion and disease prevention among diverse population

nificant).[6] Additionally, the reliability of the original buffer strip method was tested by adding each saliva sample to duplicate strips, which were found to have consistent reliability.[6] However, there is a lack of evidence in the literature on inter–examiner reliability using the various salivary tests to measure salivary characteristics. This test of reliability is especially important if the salivary characteristics are to be measured across multiple practice settings and by a variety of clinicians in a practice–based research network.

There is increasing interest in dental practice–based research networks and the types of research protocols that can be carried out in these settings.[7–10] In a given network, a research study may be carried out by upwards of 100 independent dental practitioners. Among the challenges for networks is evaluation of inter–examiner reliability for tests and measurements that are used in carrying out the protocols in practitioners' offices. Two studies described here illustrate methods that may be employed for establishing inter–examiner reliability for the protocols used in the private practices of a practice–based research network.

Northwest PRECEDENT (Practice–based REsearch Collaborative in Evidence–based DENTistry) is 1 of 3 dental practice–based research networks funded and established in 2005 by the National Institute of Dental and Craniofacial Research.[11] Member–dentists in the PRECEDENT network are drawn from the 5 state region of Washington, Oregon, Idaho, Montana and Utah. Faculty and staff at the University of Washington (UW) and the Oregon Health & Science University Schools of Dentistry have oversight and management responsibilities for the network. There are 216 fully trained member–dentists in the network including a sub–network of 44 orthodontists.

The practitioner–members suggest topics of research interest and the academic centers organize and develop protocols and materials for conducting studies in the network practices. One of the topics of primary interest to practitioners is the validity of techniques that are available for caries risk assessment. To this end, Northwest PRECEDENT is conducting Study 002: Salivary Markers in Caries Risk Assessment. This study will evaluate the contributions of historical, environmental and behavioral factors and salivary characteristics to caries risk. The primary outcome measure is caries incidence in permanent teeth over 2 years.

Prior to the implementation of tests measuring salivary characteristics in this cohort study on caries risk assessment in a practice–based setting, 2 inter–examiner reliability studies were conducted.

The objective of this report is to investigate the inter–examiner reliability of 4 salivary diagnostic tests: resting salivary pH, stimulated salivary flow rate, pH and buffering capacity in 2 populations. The first study, Reliability Study 1, was conducted on a convenience sample of dental students from the UW. The second, Reliability Study 2, was performed with patient–populations representative of the practice–based network in 4 private practices of PRECEDENT member–dentists. Reliability Study 2 also assessed the feasibility of conducting these 4 tests and an additional 2 tests (resting salivary consistency and resting salivary flow rate from labial salivary glands) by dental personnel in practice–based settings.

Methods and Materials

Protocol development established 6 salivary diagnostic tests to be used in Study 002. The salivary characteristics selected had to have evidence of potential to predict future caries.[1–3] In addition, Northwest PRECEDENT practitioners expressed that it was important that the in–office methods for salivary testing be easy–to–use and feasible to employ in a busy private practice or community clinic setting. Because they had to be easy–to–use chair–side by dental auxiliary personnel, microbiological tests were considered too cumbersome at the time to be practical, as they required special handling, a counter–top incubator and delayed reporting of results.

The salivary characteristics recorded in Study 002 are:

1. Resting salivary consistency: the appearance of the saliva is visually assessed as watery and clear or thick, sticky and frothy (Normal: watery and clear)
2. Resting salivary flow rate from labial salivary glands: the lower labial buccal mucosa is dried and the rate of secretion from the minor salivary glands is timed, up to a maximum of 90 seconds (Normal: 60 seconds or less)[12]
3. Resting salivary pH: the patient expectorates saliva (without stimulation) into a collection cup. The pH of the saliva is immediately evaluated with a pH test strip. The strip is compared to a reference chart that is provided in the data collection form (Normal: pH 6.5 to 7.5)[1] (Figure 1)
4. Stimulated salivary flow rate: the patient chews a wax pellet and expectorates into a collection cup regularly for 5 minutes. The volume of saliva is measured by reading the level of the watery component. The milliliters per minute flow rate is calculated for analysis (Normal: 1 ml per minute or greater)[2] (Figure 1)

Figure 1: Illustrations of the four salivary tests evaluated for inter–examiner reliability

Resting Salivary pH
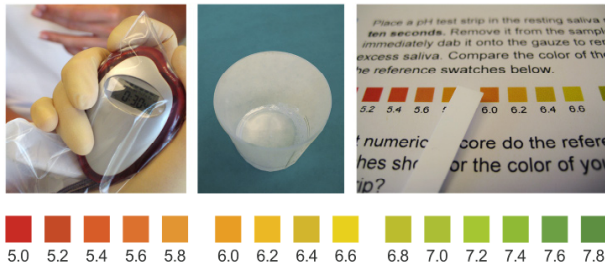
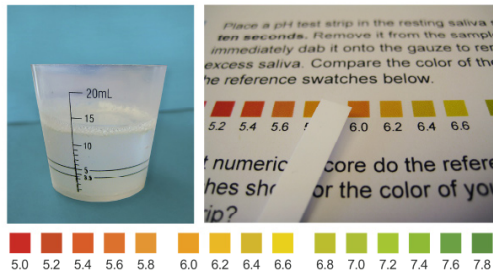Record the numerical score for the color of the test strip: [Score]

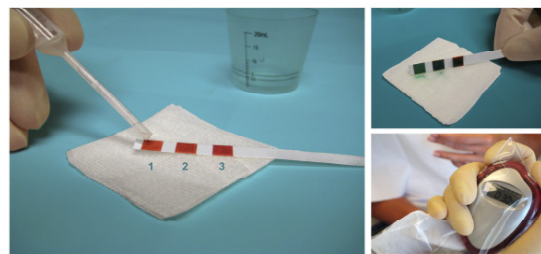## Stimulated Salivary Flow Rate

Measure the volume of saliva: [mL]

Stimulated Salivary pH

Record the numerical score for the color of the test strip: [Score]

Salivary Buffering Capacity

5. Stimulated salivary pH: the pH of the stimulated saliva sample is directly measured with a second pH test strip (Normal: pH 6.5 to 7.5)[1] (Figure 1)

6. Stimulated salivary buffering capacity: a drop of stimulated saliva is placed via pipette onto each of the 3 pads on a buffer test strip. Each pad presents a different acid challenge. The color produced on each pad by the saliva is evaluated and the corresponding verbal description on the data collection form is checked. No reference chart of colors is provided (Figure 1). The pads are individually scored from 0 to 4 (red=0, green=4). The buffer strip may have a cumulative score from 0 to 12. Examiners are not informed of the scoring code or of the significance of the colors (Normal: 10 to 12 points total per manufacturer)[13]

Salivary test kits for Study 002 containing all the necessary components for the 6 tests were developed and assembled by PRECEDENT staff. It was practical to use a commercially available buffer test strip (the Saliva–Check buffer test strip from GC Corporation, Tokyo, Japan). Additionally, timers are supplied to all practices (Figure 2).

Focus Group

After finalizing the salivary diagnostic tests to be used in Study 002, the next step developed a set of instructions for performing the tests, which were then pilot tested in a focus group setting. These instructions, which included photographs illustrating the 6 tests and reference charts for pH values, also served as the data collection instrument. Five dental assistants (not color blind) of varying levels of experience working in the dental clinics at the UW School of Dentistry were asked to participate in the focus group.

Using the data collection form with instructions for each of the salivary tests, the dental assistants individually performed the tests on a volunteer without any additional training, with the exception of the opportunity to read the instructions prior to the clinic session. The lead coordinator for Northwest PRECEDENT observed each dental assistant as salivary tests were conducted, and the coordinator evaluated the results. After the clinic session the dental assistants assembled to discuss performing the salivary tests, the feasibility and amount of time required, the instructions provided and recommendations for training. Based on their feedback, the instructions were modified and a detailed training protocol was developed.

Reliability Study 1

The purpose of this reliability study was to evaluate the degree to which dental staff, trained to perform and evaluate the results of the salivary tests, give consistent and reliable results when examining the same saliva samples. Inter–examiner reliability was evaluated for 4 of the 6 salivary tests used in Study 002: resting salivary pH, stimulated salivary flow rate (5 minute volume), pH and salivary buffering capacity (Figure 1). The resting salivary consistency test and the evaluation of resting salivary flow rate from labial salivary glands required a clinical exam that was not feasible to perform on the dental student volunteers during a class laboratory session. The UW Institutional Review Board approved the study using an oral consent process, as no identifying information was collected as part of the study and saliva samples were immediately discarded after the tests were completed.

To establish whether or not there is good inter–examiner reliability for these salivary tests, 5 UW dental assistants, different from the dental assistants who participated in the focus group, were recruited. These dental assistants were not color–blind, and were as similar as possible to the persons who would be performing the salivary tests in the dental practices of Northwest PRECEDENT. They were trained to perform and evaluate the salivary tests prior to participating in this reliability study. The training protocol involved a review of the salivary test instructions and kit components with the lead coordinator. The dental assistants then performed a practice run to administer the tests unsupervised on a volunteer. The lead coordinator clarified any questions raised when the dental assistants reported on their practice session.

During a regularly scheduled class session in the dental school simulation laboratory, 40 dental students provided salivary samples for inter–examiner reliability testing. One hour prior to the collection of salivary samples, the students were reminded that they must abstain from smoking, eating, drinking (except water), tooth brushing and using mouthwash. The students provided a resting salivary sample and a stimulated sample following the collection protocol for the salivary tests.

The 40 paired, numbered samples were immediately transported to a biomedical laboratory, where the 5 dental assistant examiners and the lead coordinator completed the 4 salivary tests on 20 pairs of samples. An incomplete block design was used to assign 20 pairs of samples to each examiner, and each sample was evaluated by 3 examiners. The examiners first evaluated all 20 assigned stim-

Figure 2: PRECEDENT salivary test kits

ulated saliva samples for flow rate (5 minute volume). After flow rate evaluation was completed, the examiners evaluated the stimulated salivary pH, buffering capacity and resting salivary pH for their assigned samples.

Reliability Study 2

The second reliability study had 2 objectives. The first was to evaluate the inter–examiner reliability of 4 of the 6 salivary tests (same tests as in Reliability Study 1) among dental assistants in practice–based settings. The resting salivary consistency test and the resting salivary flow rate from labial salivary glands could not be evaluated for inter–examiner reliability in the practice setting, as these tests were required to be performed directly on the patient, and saliva collection would change the patient's oral environment from the first to the second evaluation. The second objective was to assess the feasibility of using the 6 tests to measure salivary characteristics in a practice–based setting by evaluating the time required for completion of the tests.

Four Northwest PRECEDENT practices volunteered as sites for the inter–examiner reliability testing. It required 2 dental assistants to be available to evaluate the saliva samples back–to–back. The practice sites received study binders containing in–depth instruction on study conduct, including the Patient ID Log, phone recruitment script, Staff Log and training requirements, Manual of Procedures (MOP), salivary test instruction/data collection forms and consent forms. PRECEDENT assembled salivary test kits were provided to the sites. The lead coordinator held a conference call with the 2 salivary administrator dental assistants in each practice. The call reviewed, in detail, the

study protocol, MOP and the required training. Each dental assistant practiced administering the 6 salivary tests on an office volunteer with the dentist observing. Dentist certification of the salivary practice runs were faxed to PRECEDENT prior to starting the study.

The UW Institutional Review Board approved the study protocol using the oral consent process for patients. A random sample of 20 to 25 patients per practice was selected using an assigned systematic sampling interval (every Nth patient) based on a typical weekly patient schedule in order to recruit approximately 1 patient per day. Patients age 9 years and older who spoke English were eligible to participate.

After consent to participate in the study was obtained, it was confirmed that the patient had abstained for the previous hour from smoking, eating, drinking (except water), tooth brushing and using mouthwash. The first dental assistant (DA1) timed and performed the 6 salivary tests on the patient in the following order: resting salivary consistency, flow rate from labial glands, resting pH, stimulated salivary buffering capacity, pH and flow rate (5 minute volume). The patient's participation in the study was complete at that point. Within a maximum of 10 minutes, the second dental assistant (DA2), blind to DA1's results, evaluated the same samples to determine the inter–examiner reliability for 4 of the 6 tests in the following order: measurement of stimulated salivary flow rate (5 minute volume), buffering capacity, pH and resting salivary pH. The saliva samples were then discarded. The dental assistants' roles were not interchangeable.

Statistical Analysis

Separate analyses of reliability of the salivary diagnostic tests were conducted for the 2 studies. Descriptive statistics (mean and standard deviation (SD)) were calculated for each of the tests. Reliability was measured for each test using the intraclass correlation coefficient (ICC), which is defined as the correlation between 2 test results obtained by 2 different examiners for the same patient. The ICC and 95% confidence interval (CI) was estimated using an analysis of variance method.[14] For Reliability Study 1, a balanced incomplete block design was used with both examiner and subject as random effects. For Reliability Study 2, examiner and subject were nested within practice, and practice, examiner and subject were random effects. Statistical analyses were conducted using Stata version 8 and were independently validated.[15] The following scheme was used for interpretation of ICC values: 1) ICC≤0.2 (no or very low agreement), 0.21 to 0.40 (low agreement), 0.41 to 0.60 (moderate agreement), 0.61 to 0.80 (high agreement) and 0.81 to 1.00 (excellent agreement).[16]

Results

Reliability Study 1

Identifying information for the participating dental student subjects, such as age and gender, was not collected. However, the mean age of the entire class of students was 26 years old and 38% of the class was female. The mean stimulated salivary flow rate (calculated from the 5 minute volume) was 1.78 ml per minute, with a mean stimulated salivary pH of 7.61. The resting salivary pH mean was 7.28. Buffer pad 1 had a mean score of 3.83, buffer pad 2 had a mean of 3.63 and buffer pad 3, the strongest acid challenge, had a mean of 1.67. This resulted in an overall buffer capacity mean score of 9.13 (Table I).

The assessment of stimulated salivary flow rate demonstrated excellent inter–examiner reliability (ICC=0.96). The resting salivary pH showed high inter–examiner reliability (0.76), while the stimulated salivary pH had a very low ICC (0.08). The stimulated salivary buffering capacity test, with the 3 test pad scores summed, had moderate inter–examiner reliability (ICC=0.43), with very low reliability of the first and second test pad challenges (ICC=0.02 and 0.20) and moderate reliability of the third acid challenge (ICC=0.46) (Table I).

Reliability Study 2

Among 4 Northwest PRECEDENT dental practices, 85 patients were recruited for inter–examiner reliability testing. The number of patients in each age category, as designated for Study 002, were as follows: age 9 to 17 (n=23), age 18 to 64 (n=45) and age 65+ (n=17). DA1 conducted the 6 salivary tests in a mean time of 13 minutes per patient, with a range of 8 to 17 minutes (SD=2.2 minutes). The average time between DA1 and DA2 conducting their respective salivary tests was 5.5 minutes (SD=9, median=3, interquartile range=1 to 7 minutes).

The first 2 salivary diagnostic tests, resting salivary consistency and resting salivary flow rate from labial salivary glands were conducted by DA1 only. Resting salivary consistency was thick and frothy in 18% of the participants. The mean time for the labial glands to produce saliva was 41 seconds (SD=25, median=35, interquartile range=22 to 53 seconds). A cut–off of 90 seconds was used for the

Table I: Summary of the salivary tests for Reliability Study 1

Salivary Test	Mean (SD)	ICC (95% CI)
Rate (ml / minute)	1.78 (0.74)	0.96 (0.93–0.98)
Resting pH	7.28 (0.33)	0.76 (0.62–0.85)
Stimulated pH	7.61 (0.17)	0.08 (0.00–0.29)
Buffer pad 1	3.83 (0.49)	0.02 (0.00–0.23)
Buffer pad 2	3.63 (0.55)	0.20 (0.00–0.41)
Buffer pad 3	1.67 (1.49)	0.46 (0.26–0.64)
Buffer capacity	9.13 (1.71)	0.43 (0.22–0.61)

Table II: Summary of the salivary tests for Reliability Study 2

Salivary Test	Mean (SD)	ICC (95% CI)
Rate (ml / minute)	1.29 (0.78)	0.94 (0.91–0.96)
Resting pH	6.67 (0.46)	0.82 (0.74–0.88)
Stimulated pH	7.42 (0.27)	0.80 (0.71–0.87)
Buffer pad 1	3.64 (0.66)	0.23 (0.01–0.42)
Buffer pad 2	2.67 (1.53)	0.46 (0.25–0.60)
Buffer pad 3	0.99 (1.45)	0.27 (0.05–0.46)
Buffer capacity	7.34 (2.74)	0.55 (0.37–0.68)

12 participants who did not produce labial gland saliva within the test time limit.

For each patient's saliva samples, 4 salivary characteristics were evaluated and recorded by both DA1 and DA2. The mean stimulated salivary flow rate (calculated from the 5 minute volume) in this patient population was 1.29 ml per minute, and the mean pH of the stimulated saliva was 7.42. The mean resting salivary pH was 6.67, and the stimulated salivary buffering capacity mean score was 7.34 (Table II).

Assessment of stimulated salivary flow rate (ICC=0.94) and resting salivary pH (ICC=0.82) demonstrated excellent inter–examiner reliability, while stimulated salivary pH (ICC=0.80) showed high inter–examiner reliability. The stimulated salivary buffering capacity test had moderate reliability (ICC=0.55) (Table II).

Discussion

Two reliability studies assessed 4 salivary tests in a convenience sample of dental students and a sample of dental patients from 4 representative PRECEDENT practices. The inter–examiner reliability of stimulated salivary flow rate and resting salivary pH were considered very good in both studies, but the stimulated salivary pH and buffering capacity were not acceptable in 1 study.

In Reliability Study 1, the small variation in stimulated salivary pH and buffering capacity among the dental students may have artificially made the reliability appear low. For this reason, it was determined that further inter–examiner reliability testing enrolling a wider variety of subjects with a potentially greater variation in salivary characteristics was needed. The follow–up study was conducted in a sample of PRECEDENT practitioner offices with a patient population representative of the dental practices where these tests are intended to be used.

In Reliability Study 2, the reliability of stimulated salivary flow rate and resting salivary pH were consistent with Reliability Study 1, while the stimulated salivary pH and buffering capacity tests had a better performance. The stimulated salivary pH presented a high agreement. While the reliability of the individual pads with different acid challenges for the buffering capacity test was low or at the low end of moderate (pad 2), the overall buffering capacity test score had improved moderate reliability.

The lower reliability of the buffer capacity may have been due to the ambiguity in color interpretation of the buffering strip test pads with the 3 different acid challenges. A reference chart is not available as a visual aid for evaluation of the colors as it is for the pH tests. This result correlates with the evaluation of the earlier chair–side tube method of estimating salivary buffering capacity, where the method was precise only within the crude criteria of low, intermediate and high buffer capacity compared to the standard electrometric method.[17] This tube method evolved into the buffer strip method used today. Other buffer strips continue to use the low, intermediate and high criteria, rather than attempting to measure buffering capacity at differing levels of acid challenge. Those buffer strips that use the low, intermediate and high criteria provide a color reference chart for evaluating the results. A similar guide for color interpretation could improve the precision of evaluating the results from the buffer test strip employed in these studies.

The 4 dental practices conducting this study found the salivary diagnostic tests easy to incorporate into the routine of the day. Dental personnel were quickly and easily trained to perform the tests and patients were accepting of the procedures. The greatest challenge was for patients to remember not to eat, drink (except water), brush their teeth or use mouthwash within the hour prior to their appointment. In some cases, the office staff contacted patients an hour before the appointment to provide a reminder. New technology in the form of automated

electronic texting of appointment reminders has the ability to minimize this problem in the future.

Practice–based research networks have been established to assess their ability to influence the evidence–base of dental practice.[11] There is a question whether a practice–based network can or should attempt to perform strict calibration and reliability testing of measures in their studies, or whether the variability and lack of calibration can be tolerated because it is off–set by the large numbers of evaluations that can be made in a network.[8,18–20] Some practice–based research network participants advocate for less emphasis on measurement calibration and reliability because that reflects the "real–life" setting of dental practice. To do otherwise might not be palatable to practitioners.[9]

If practice–based networks are to provide the best possible evidence–base for the practice of dentistry, then sound scientific methodology is an appropriate goal when conducting research. In a network setting, where multiple examiners are collecting data and taking measurements, it is possible to establish a means to assess the inter–examiner reliability of tests and measurements used. The process described here was the basis for a training protocol whereby all examiners had the same detailed level of instruction and practice in performing procedures prior to study initiation. The training protocol having been established, inter–examiner reliability was evaluated for the salivary tests used. When equivocal results were obtained, further testing was determined to be necessary, along with possible adjustments to the training and procedures employed. In this case, further testing established adequate inter–examiner reliability levels for use of these tests and training protocol in the study on Salivary Markers in Caries Risk Assessment, Study 002 of the Northwest PRECEDENT network.

Conclusion

Patients' salivary characteristics may provide information about future caries risk. Prior to implementation of tests measuring salivary characteristics in a cohort study on caries risk assessment in a practice–based research network, 2 inter–examiner reliability studies were conducted, the first on a convenience population of dental students, and the second with patient–populations representative of the practice–based network. These studies demonstrated acceptable inter–examiner reliability for 4 salivary diagnostic tests: stimulated salivary flow, pH, buffering capacity and resting salivary pH. In addition, it showed that it is feasible to perform the 6 tests measuring salivary characteristics in practice–based settings in terms of time, personnel required and patient acceptance.

Marilynn Rothen, BS, RDH, is Lead Regional Coordinator for Northwest PRECEDENT; Joana Cunha–Cruz DDS, PhD, is research assistant professor, Department of Dental Public Health Sciences, and epidemiologist for Northwest PRECEDENT; Lloyd Mancl, PhD, is research associate professor, Department of Dental Public Health Sciences and biostatistician for Northwest PRECEDENT; Brian Leroux is professor, Biostatistics and Department of Dental Public Health Sciences, and principal investigator for the Northwest PRECEDENT Data Coordinating Center; Brooke Latzke Davis, MPH, is a research associate, Seattle Children's Research Institute; Justin Coyne was Network Media and Information Management Specialist, Northwest PRECEDENT; and Joel Berg, DDS, MS, is Professor & Chair, Department of Pediatric Dentistry, Associate Dean for Hospital Affairs and co–investigator for Northwest PRECEDENT; all are at the University of Washington School of Dentistry, Seattle, WA. Jane Gillette, DDS, is a practitioner–member of Northwest PRECEDENT located in Bozeman, Montana.

Acknowledgments

References

1. Tenovuo J. Salivary parameters of relevance for assessing caries activity in individuals and populations. *Community Dent Oral Epidemiol*. 1997;25(1):82–86.

2. Leone CW, Oppenheim FG. Physical and chemical aspects of saliva as indicators of risk for dental caries in humans. *J Dent Educ*. 2001;65(10):1054–1062.

3. Fontana M, Zero DT. Assessing patients' caries risk. *J Am Dent Assoc*. 2006;137(9):1231–1239.

4. Ericsson Y. Clinical investigation of the salivary buffering action. *Acta Odontol Scand*. 1959;17(2):131–165.

5. Frostell G. A colourimetric screening test for evaluation of the buffer capacity of saliva. *Swed Dent J*. 1980;4(3):81–86.

6. Ericson D, Bratthall D. Simplified method to estimate salivary buffer capacity. *Scand J Dent Res*. 1989;97(5):405–407.

7. Wotman S, Lalumandier J, Nelson S, Stange K. Implication for dental education of a dental school–initiated practice research network. *J Dent Educ*. 2001;65(8):751–759.

8. Burke FJ. Evaluating restorative materials and procedures in dental practice. *Adv Dent Res*. 2005;18(3):46–49.

9. Mjör IA. Practice–based dental research. *J Oral Rehabil*. 2007;34(12):913–920.

10. Chattopadhyay A, Arevalo O, Sohn W. Understanding measurement of dental diseases and research participation in practice set–up. *Dent Clin North Am*. 2008:52(2):367–386.

11. Kuska B. NIDCR Awards Grants for New Practice–Based Initiative. NIDCR [Internet]. 2005 [cited 2009 Aug 13]. Available from: http://www.nidcr.nih.gov/Research/ResearchResults/NewsReleases/ArchivedNewsReleases/NRY2005/PR03312005.htm

12. Walsh LJ. Preventive dentistry for the general dental practitioner. *Aust Dent J*. 2000;45(2):76–82.

13. Saliva Check Technical Manual. Tokyo: GC Corporation. 2002.

14. Rosner B. Fundamentals of Biostatistics. 7th ed. Belmont: Thomson–Brooks/Cole. 2010. 568–571 p.

15. StataCorp. Stata Statistical Software: Release 8. College Station, TX: StataCorp LP. 2003.

16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174.

17. Wikner S, Nedlich U. A clinical evaluation of the ability of the Dentobuff method to estimate buffer capacity of saliva. *Swed Dent J*. 1985;9(2):45–47.

18. Mjör IA. Controlled clinical trials and practice–based research in dentistry. *J Dent Res*. 2008;87(7):605.

19. Derouen TA, Ferracane J, Mjör IA. Controlled clinical trials and practice–based research in dentistry. *J Dent Res*. 2008;87(9):800.

20. Finkelman RD, Polson AM. Controlled clinical trials and practice–based research in dentistry. *J Dent Res*. 2008;87(12):e2.