A Closer Look at Diagnosis in Clinical Dental Practice: Part 4. Effectiveness of Nonradiographic Diagnostic Procedures and Devices in Dental Practice

• **Gerardo Maupomé**, PhD •
• **Iain A. Pretty**, BDS(Hons), MSc, PhD •

Abstract

This article, the fourth in a series, examines nonradiographic procedures and devices such as standard clinical and visual examination, apex locators, vitality testers and colour shade guides in light of the tools described in the first 2 articles in the series. A variety of nonradiographic indices and scales are used in detecting periodontal disease and monitoring and assessing its treatment. The reliability of these diagnostic procedures directly affects treatment success, decisions to initiate more aggressive clinical interventions, and the ability to make an informed prognosis about the course of the disease. However, in many instances, the dependability of the measurements remains to be established.

MeSH Key Words: *decision support techniques; predictive value of tests; risk assessment methods*

This article has been peer reviewed.

The first 2 articles of this series[1,2] introduced some of the basic concepts used in assessing diagnostic accuracy: reliability, validity, sensitivity and specificity, as well as predictive values and receiver operating characteristic (ROC) analysis. Part 3 of the series[3] used the tools described in those first 2 articles to examine the most common dental diagnostic procedures involving radiography. Part 4 does the same for common nonradiographic procedures and devices such as standard clinical and visual examinations, apex locators, vitality testers and colour shade guides.

A glossary, with concise definitions of terms, is available for the entire series (see **Appendix 1**, Glossary of epidemiology terms, at http://www.cda-adc.ca/jcda/vol-70/issue-4/251.html).

Nonradiographic Periodontal Measurements

Probing Pockets for Attachment Level

Probing pockets for attachment level is a mainstay of periodontal practice. One study investigated the reliability of periodontal probing by contrasting the most commonly used method — manual probing — with electronic probing.[4] The researchers compared data obtained with a regular dental probe with those generated by an automated force-controlled probe, testing a total of 1,128 sites in 15 patients undergoing periodontal maintenance. Probing depth and relative attachment level were recorded to the nearest 0.5 mm with both instruments on 2 separate visits. At shallow sites (3 mm or less), reproducibility of probing depth with the manual probe was 59.1% for exact agreement and 98.6% for variation within 1.0 mm; values for the electronic probe were 41.3% and 91.5%, respectively. Reproducibility of probing depth measurement was poorer at deeper sites, with an exact match of only 33% for manual probing and 32% for the electronic probe. The researchers found corresponding, but lower, reliability for attachment level readings. Overall, there was no reliability advantage with the electronic probe.

In another study assessing accuracy with several probes of each type (Marquis, William and EN-15 manual probes and Florida Pocket Probe, Florida Disk Probe and

Table 1 **Diagnostic devices tested**

Diagnostic device	Manufacturer
Probes	
Manual probes	
Marquis	Ash Instruments, Dentsply, Woodbridge, Ont.
William	Ash Instruments, Dentsply
EN-15	Ash Instruments, Dentsply
Electronic probes	
Florida Pocket Probe	Florida Probe Corp, Gainsville, Fla.
Florida Disk Probe	Florida Probe Corp
Peri Probe	Samhall Pile Dental, Malmö, Sweden
Apex locators	
Apex Finder AFA (All Fluids Allowed) model 7005	Analytic Endodontics, Orange, Calif.
Apex-Finder	Analytic Endodontics
Neosono Ultima EZ	Satelec, Toronto, Ont.
Apit 2	Osada Inc, Tokyo, Japan
Odometer	L. GooF A/S, Hörsholm, Denmark
Endocater	Hygienic Corp., Akron, Ohio
Shade guides	
Vita Lumin Vacuum	Vident, Brea, Calif.
Colortron II	X-Rite, Grandville, Mich.
Vitapan 3-D Master	Vident
Vita EasyShade	Vident
Shade-Eye	Shofu, San Marcos, Calif.

Peri Probe electronic probes; **Table 1**), the electronic probes offered significantly greater accuracy (percentage correct) than manual and conventional probes.[5] Interestingly, experience in probing was an important factor influencing accuracy but not reliability within this study. These results confirmed the findings of Wang and others,[4] who reported similar reproducibility for all probes of the same type. The William and Florida probes performed best overall.

Observing Bleeding on Probing

A common test of gingival health — the observation of bleeding on probing — constitutes an important part of many periodontal examinations. In a study of 41 patients that was designed to determine the reliability of bleeding on probing as an indicator of gingival status,[6] this method had a sensitivity of only 29% but a specificity of 88%. The positive predictive value was 6%, and the negative predictive value 98%. These data suggest that absence of bleeding is a good indicator of gingival health. This finding was supported by another study, which found that bleeding on probing was one of the most reliable predictive indicators of further attachment loss over a 42-month period.[7]

Measuring Furcation

The periodontal involvement of tooth furcations is used as a marker of more advanced disease and is often considered a reliable indicator that more aggressive therapies are appropriate and that prognosis for the affected site is poor.[8] Interestingly, the anatomic location of the furcal involvement seems to affect the reproducibility of assessments. In a study examining furcation measurements in 100 molars in 25 patients,[9] the level of agreement, as indicated by kappa coefficients, was excellent for buccal, lingual and mesiolingual furcations (0.77 to 0.94) but only moderate for distolingual lesions (0.70). When assessing the reliability of measurement of 125 furcations in 60 molars with 2 different probe types, the intraclass reliability was recorded as 0.67.[10]

Using Periodontal Measurements as Predictors of Disease

In some cases, periodontal measurements obtained through various diagnostic procedures are combined to evaluate overall treatment needs. This combination approach has itself been evaluated by determining the number of diagnostic procedures that must be performed (and their costs) before relative saturation of information is reached (at which point another test yields no new information). Loesche and others[11] factored in tooth type, furcation involvement, bleeding on probing, attachment level, probing depth, mobility, and benzoyl-DL-arginine-naphthylamide (BANA) test scores to determine the validity of such measures in predicting periodontal disease severe enough to require surgical intervention. Using an ROC method, the researchers determined that the combination of these measures had a sensitivity of 76% and a specificity of 75%. The practical question that remains is whether the clinical time required to attain such high levels of sensitivity and specificity can be devoted to individual patients in everyday clinical practice.

Nonradiographic Endodontic Devices

Apex Locators

Apex locators are becoming increasingly common in both general and specialized practice. The ability to quickly check canal length without recourse to radiography is of benefit to both patient and operator. Several studies have looked at the accuracy and reliability of these devices. Using a variety of in vitro models (including wet and dry canals), DeMoor and others[12] measured the differences between canal length indicated by 4 apex locators — Apex Finder AFA model 7005, Apex-Finder, Neosono Ultima EZ and Apit 2 (**Table 1**) — and actual canal length. Only the Apex-Finder was unreliable, providing measurements that were more than 0.5 mm from the apical foramen. When a ranking system based on a precision level of ± 0.1 mm from the apical foramen was used, the Apex Finder AFA model 7005 was the most accurate device.

The use of apex locators in clinical settings has been studied, and some results suggest that electronic readings of canal length are more reliable than those obtained from radiographs.[13,14] This application may have repercussions in the choice of diagnostic procedures, since using an electronic apex locator might enable the clinician to reduce the number of radiographs required during endodontic procedures. In an in vivo test of 2 apex locators, measurements taken within the two 0.5-mm intervals closest to the apical constriction were 84.8% (Odometer) and 89.6% (Endocater) of the working length determined after extraction.[15]

Figure 1: *An 11-year old boy after traumatic fracture of the central incisor (a) and a 32-year-old woman reporting discoloration of her upper left central incisor (b). Using a variety of tests, dentists can determine if the teeth shown in these images are vital. However, even with a positive test result, how sure can we be that the teeth are in fact vital? Diagnostic test values suggest that the probability of a sensitive reaction (representing vital pulp) would be 90% with a cold test, 83% with a heat test and 84% with electronic pulp testing.[16]*

Methods of Testing Vitality

Several devices and techniques enable the clinician to determine if a tooth is vital or not. Commonly used diagnostic procedures include a pain history, testing with ethyl chloride (cold stimulus in thermal testing), application of heat (hot stimulus in thermal testing), electronic pulp testing, assessment of tooth colour and use of radiographs to detect periapical radiolucencies (see **Fig. 1**). An investigation has been conducted to assess the diagnostic effectiveness of a new pulpal test,[17] and while that particular test is not considered here, the study also estimated sensitivity values for the more traditional techniques: 92% for ethyl chloride, 36% for radiographic views and 16% for pain history. From these data, it can be readily appreciated that the operating characteristic values for the 2 latter procedures were quite poor, and a substantial proportion of suspect teeth might not be identified as nonvital if these methods were the only means of assessing vitality status.

A separate study provides a useful approach to evaluating diagnostic procedures for determination of vitality. Petersson and others[16] examined a range of accuracy data for ethyl chloride, hot gutta-percha and an electronic pulp tester applied to teeth with normal radiographic appearance; the gold standard was direct clinical inspection of the pulp. Their findings offer useful comparators for this discussion. Sensitivity was 0.83 for the cold test, 0.86 for the heat test and 0.72 for the electrical test; specificity was 0.93, 0.41 and 0.93, respectively. Positive predictive value was 0.89 for the cold test, 0.48 for the heat test and 0.88 for the electrical test; negative predictive value was 0.90, 0.83 and 0.84, respectively. The results concerning sensitivity, specificity and negative and positive predictive values can be interpreted as follows: the probability that a nonsensitive reaction (i.e., a positive test result) represented a necrotic pulp was 89% when the cold test was used, 48% when the heat test was used and 88% when the electrical test was used. The probability that a sensitive reaction (i.e., a negative test result) represented a vital pulp was 90% when the cold test was used, 83% when the heat test was used and 84% when the electrical test was used. Although these values are generally high, in about 1 instance out of 10 a negative result will be false, and the patient may undergo unnecessary root canal therapy. A good rule of thumb may be to use more than one diagnostic procedure, to increase the degree of diagnostic certainty, but the findings of Petersson and others[16] indicate that the traditional devices (e.g., ethyl chloride) sometimes perform better than their newer, more expensive counterparts.

Shade Guides for Assessment of Tooth Colour

The placement of tooth-coloured restorations and crowns is an important component of restorative and rehabilitative dentistry. The ability to accurately and reliably assess colour shades in this context is crucial to a successful clinical outcome and to fulfilling patient expectations. Typically, dentists use a shade guide to select either a restorative material of equivalent shade or to communicate instructions to a dental laboratory. Because this is a highly subjective process, researchers have developed a range of electronic devices to assess colour. These new devices, such as Vident's EasyShade and Shofu's Shade-Eye (**Table 1**), will be discussed in detail in the sixth article in this series.

Here, we focus on the reliability and accuracy of more traditional shade guides. Comparing the performance of clinicians using a shade guide is complex, since there is no obvious way to establish a gold standard. To illustrate such performance, we briefly turn to a study comparing a visual shade guide (Vita Lumin Vacuum Shade Guide) with a new electronic device (Colortron II) (**Table 1**).[18] Using 16 coloured tabs, 31 observers made repeated colour assessments. The observers averaged 7.7 correct matches (48% of total); the repeatability index for this assessment was rated as fair (r = 0.60). The study's authors concluded that shade determination by visual means was inconsistent but that the new colorimeter performed only marginally better. More recently, a new shade guide has been introduced, the Vitapan 3-D Master (**Table 1**), which is claimed to offer a more accurate colour-matching system. In a study comparing, among other systems, the traditional Vita Lumin Vacuum guide with the Vitapan 3-D Master,[19] the number of clinically acceptable colour matches was 46% for the traditional guide and 56% for the 3-D guide, but this difference was not statistically significant.[19] One source of concern was the low number of matches that would be considered clinically acceptable according to U.S. Public Health Service (USPHS) criteria proposed by Ryge.[20] Overall, there is a need for more accurate and reliable colour-matching systems.

Discussion

The studies discussed in the third[3] and fourth articles of this series have showcased the strengths of some commonly used diagnostic procedures, as indicated by their accuracy, validity and reliability. The most striking feature recurs across all of these applications: the clinical value of a specific diagnostic procedure may or may not be associated with its operating characteristics. As such, this information may be ascribed greater or lesser weight in the realm of clinical decision-making, often for nonclinical reasons.[21] Discrepancies among the results of the different diagnostic procedures available in dental practice should not be a reason to dismiss the information thus supplied. Rather, information obtained from multiple diagnostic procedures should be objectively compared under certain guiding principles.[22]

The identification of appropriate procedures to prevent, diagnose and treat dental disease is challenging for several reasons. First, although we know a lot about these diseases — in particular, dental caries and periodontal diseases — much remains unknown. Information is imperfect, yet dental clinicians are expected to make decisions about individual patients every day, decisions that will be based at least partially on probabilistic, rather than definitive, data. A frequently mentioned example is the poorly understood array of dental problems among the expanding numbers of elderly people, who are living longer and retaining functional dentition well into old age.

Second, patients differ from one another in clinically important ways — in clinical presentation, in the courses of disease and of health, in their ability to adhere to preventive and treatment regimens, in the values to which they subscribe and in their preferences for treatments and outcomes. Uncertainty abounds about their risk of dental disease, about diagnostic and prognostic information, about the efficacy and effectiveness of many management and treatment alternatives, and about the outcomes associated with various clinical strategies. It is unlikely that dental diagnostic procedures will ever be able to address all of these important, and clinically relevant, questions.

Third, the wealth of evidence that informs decisions about diagnosis and management of dental diseases is continuously evolving. New diagnostic procedures are constantly being introduced as technologies expand. Indeed, in some cases, current knowledge provides only a partial understanding of specific disease problems, leaving the clinician to rely on subjective clinical expertise, which may or may not result in appropriate clinical management.

Competing goals and multiple perspectives often influence clinical decisions. Patients and their caregivers may have values and preferences for treatment options and outcomes that differ from those of practitioners. Health care delivery models and clinical systems have priorities, policies and funding limitations that curtail the availability of certain clinical measures.

Given the features and limitations of typical clinical decision making, a good understanding of the strength of various diagnostic procedures, based on an objective appraisal of their operating characteristics, is invaluable to the dental professional. Placing specific values (objectively derived from operating characteristics) on individual diagnostic procedures should theoretically allow a more precise evaluation of what can be expected from procedures when applied in the clinical setting. Ideally, such evaluation would go beyond less desirable influences, such as degree of familiarity with certain devices, industrial marketing efforts or the novelty of technological innovation per se.

The final 2 articles of the series will examine some of the most recent innovations in diagnostic procedures and will succinctly assess their potential to become mainstream tools for the individual dental clinician. ◆

Dr. Maupomé is investigator, Center for Health Research, Portland, Oregon; assistant adjunct professor, University of California at San Francisco, San Francisco, California; and clinical professor, University of British Columbia, Vancouver.

Dr. Pretty is lecturer and research fellow in prosthodontics, The University of Manchester, Manchester, UK.

Correspondence to: *Dr. Iain A. Pretty, Unit of Prosthodontics, Department of Restorative Dentistry, University Dental Hospital of Manchester, Higher Cambridge St., Manchester, M15 6FH, England. E-mail: iain.pretty@man.ac.uk.*

The authors have no declared financial interests in any company manufacturing the types of products mentioned in this article.

References

1. Pretty IA, Maupome G. A closer look at diagnosis in clinical dental practice: Part 1. Reliability, validity, specificity and sensitivity of diagnostic procedures. *J Can Dent Assoc* 2004; 70(4):251–5.

2. Pretty IA, Maupome G. A closer look at diagnosis in clinical dental practice: Part 2. Using predictive values and receiver operating characteristics in assessing diagnostic accuracy. *J Can Dent Assoc* 2004; 70(5):313–6.

3. Pretty IA, Maupome G. A closer look at diagnosis in clinical dental practice: Part 3. Effectiveness of radiographic diagnostic procedures. *J Can Dent Assoc* 2004; 70(6):388–94.

4. Wang SF, Leknes KN, Zimmerman GJ, Sigurdsson TJ, Wikesjo UM, Selvig KA. Intra- and inter-examiner reproducibility in constant force probing. *J Clin Periodontol* 1995; 22(12):918–22.

5. Samuel ED, Griffiths GS, Petrie A. In vitro accuracy and reproducibility of automated and conventional periodontal probes. *J Clin Periodontol* 1997; 24(3):340–5.

6. Lang NP, Adler R, Joss A, Nyman S. Absence of bleeding on probing. An indicator of periodontal stability. *J Clin Periodontol* 1990; 17(10):714–21.

7. Claffey N, Nylund K, Kiger R, Garrett S, Egelberg J. Diagnostic predictability of scores of plaque, bleeding, suppuration and probing depth for probing attachment loss. 3 1/2 years of observation following initial periodontal therapy. *J Clin Periodontol* 1990; 17(2):108–14.

8. Pihlstrom BL. Periodontal risk assessment, diagnosis and treatment planning. *Periodontol 2000* 2001; 25:37–58.

9. Eickholz P, Staehle HJ. The reliability of furcation measurements. *J Clin Periodontol* 1994; 21(9):611–4.

10. Eickholz P, Steinbrenner H, Lenhard M, Marquardt M, Holle R. Interexaminer reliability of the assessment of clinical furcation parameters as related to different probes. *Eur J Oral Sci* 1999; 107(10):2–8.

11. Loesche WJ, Taylor G, Giordano J, Hutchinson R, Rau CF, Chen YM, and other. A logistic regression model for the decision to perform access surgery. *J Clin Periodontol* 1997; 24(3):171–9.

12. De Moor RJ, Hommez GM, Martens LC, De Boever JG. Accuracy of four electronic apex locators: an in vitro evaluation. *Endod Dent Traumatol* 1999; 15(2):77–82.

13. Fouad AF, Reid LC. Effect of using electronic apex locators on selected endodontic treatment parameters. *J Endod* 2000; 26(6):364–7.

14. Ounsi HF, Haddad G. In vitro evaluation of the reliability of the Endex electronic apex locator. *J Endod* 1998; 24(2):120–1.

15. Pallares A, Faus V. An in vivo comparative study of two apex locators. *J Endod* 1994; 20(2):576–9.

16. Petersson K, Soderstrom C, Kiani-Anaraki M, Levy G. Evaluation of the ability of thermal and electrical tests to register pulp vitality. *Endod Dent Traumatol* 1999; 15(3):127–31.

17. Evans D, Reid J, Strang R, Stirrups D. A comparison of laser Doppler flowmetry with other methods of assessing the vitality of traumatised anterior teeth. *Endod Dent Traumatol* 1999; 15(6):284–90.

18. Okubo SR, Kanawati A, Richards MW, Childress S. Evaluation of visual and instrument shade matching. *J Prosthet Dent* 1998; 80(6):642–8.

19. Wee AG, Kang EY, Johnston WM, Seghi RR. Evaluating porcelain color match of different porcelain shade-matching systems. *J Esthet Dent* 2000; 12(5):271–80.

20. Ryge G. Clinical critera. *Int Dent J* 1980; 30(4):347–58.

21. Maupome G, Sheiham A. Explanatory models in the interpretations of clinical features of dental patients within a university dental education setting. *Eur J Dent Educ* 2002; 6(1)2–8.

22. White BA, Maupome G. Clinical decision-making for dental caries management. *J Dent Educ* 2001; 65(10):1121–5.