Dentist Reliability in Classifying Disease Risk and Reason for Treatment

James D. Bader, DDS, MPH; B. Alex White, DDS, DrPH; Oralee Olsen RDH, BS; Daniel A. Shugars, DDS, PhD

Abstract

Objectives: *The reliability of practicing dentists' classifications of patients' caries risk and periodontal disease risk and reason for treatment for individual teeth were determined. The risk classification protocols had been in use in a group practice for more than a year, and the reason-for-treatment protocol had been introduced six months previously.* **Methods:** *Eight dentists' classifications for caries (n=66) and periodontal disease risk (n=66), and six dentists' classifications for reason for treatment (n=73) were compared to those of a nominal standard examiner. Reliability was expressed as percent agreement and kappa values.* ***Results:*** *Percent agreement was 76 percent, 83 percent, and 74 percent for caries, periodontal disease, and reason for treatment, respectively, with kappa values of 0.56, 0.70, and 0.69.* ***Conclusions:*** *Dentists can attain reasonable levels of reliability using simple classification protocols with little formal training, although misclassification may be problematic for specific administrative or research-related purposes.[J Public Health Dent 1999;59(3):158-61]*

Key Words: Risk assessment; reliability; diagnosis, oral; dental records; decision making.

Categorizing patients by their risk of caries and periodontal disease has been advocated as an initial step in determining appropriate preventive treatment interventions (1). Several caries risk classification schemes designed for use in daily practice have been described recently, all of which incorporate one or more subjective criteria (2-4). In contrast, formal risk classification schemes for periodontal disease have not been described widely, although an understanding of risk factors for disease and disease progression is maturing (5,6). Because classification schemes are intended to help guide prevention and treatment-related decisions, their reliability and predictive validity will have implications for both cost of care and patient disease outcomes. To date, no information has been reported describing either characteristic for caries and periodontal disease risk classification schemes applied in clinical practice. Yet dentists' agreement on visual and radiographic diagnoses of caries is itself substantially less than perfect (7), and periodontal diagnoses and risk of progression classifications require agreement at several decision levels (5). Thus, the reliability of dentists' risk classifications for these two diseases merits examination at the beginning of what is expected to be an era of adoption of this developing technology.

Although currently not widely advocated, it is possible that diagnostic codes or some standardized classification scheme of reason-for-treatment also will become a routine element of the treatment record in the future. Dentists now do not routinely associate specific formal diagnoses with treatment procedures they provide; nevertheless, such information is appealing from both research and care management perspectives (8). In anticipation of the need to include diagnostic information in the treatment record for legal and/or administrative reasons, the American Dental Association will introduce a set of diagnostic codes in 1999. Because the reliability of dentists' assignment of diagnoses or reason for treatment associated with individual dental procedures will be subject to the same sources of variation as are dentists' risk classifications, an examination of the reliability of these assignments also is warranted. We report an initial assessment of the reliability of the dentists' classifications of caries and periodontal risk and reasons for treatment.

Methods

Protocols for classification of caries and periodontal risk were introduced in late 1996 into a large group practice operating 13 dental clinics. The classification protocols had low-risk, moderate-risk, and high-risk categories. For caries, the principal criterion for the low-risk category was "no new caries in prior 12 months"; the moderate-risk criterion was "new carious lesions in prior 12 months (not broken or defective restorations)"; and the high-risk criterion was "extensive or persistent caries, especially smooth surface, in prior 12 months." Modifying factors, intended to be considered in formulating treatment plans, but also applied in assigning patients to risk categories, were listed as poor family dental health, xerostomia, physical and mental handicaps, orthodontic appliances, fluoridation status, and gingival recession. For periodontal disease, the low-risk category criteria were "no pocket probing depths >4 mm, bone levels within 2 mm of CEJ on well-angled BWs." For moderate

Send correspondence to Dr. Bader, Sheps Center for Health Services Research, University of North Carolina, CB #7590, Chapel Hill, NC 27599-7590. E-mail: jim_bader@unc.edu. Web site: http://www.shepscenter.unc.edu. Dr. Shugars is with the Sheps Center for Health Services Research; Dr. White and Ms. Olsen are with Kaiser Permanente Center for Health Research, Portland, OR. Reprints will not be available. Supported by the Agency for Health Care Policy and Research, grant #I IS09453. This information was presented as a poster March 11, 1999, at the annual session of the International Association of Dental Research, Vancouver, BC, Canada. Manuscript received: 2/9/99; returned to authors for revision: 3/22/99; accepted for publication: 6/29/99.

risk, the criteria were "5–6 mm pocket probing depths (excluding third molars), bone levels 2–4 mm from CEJ on well-angled BWs." High-risk criteria were ">6 mm pocket probing depths (excluding third molars), bone levels >4 mm from CEJ on well-angled BWs." Modifying factors were receipt of periodontal treatment within the previous three years, smoking, compromised immune system, diabetes, use of medications with oral effects, family history of periodontal disease, chemotherapy or radiation therapy, age, physical and mental handicaps, and prostheses. Dentists received a sheet that listed the criteria as well as the modifying factors to be considered in formulating a preventive treatment plan. Dentists also participated in a single one-hour training session at the time the classification systems were introduced. No additional training was offered, although dentists were urged periodically to determine classifications during all new and returning patient examinations; this aspect of their performance was part of their year-end evaluations.

In late 1997 a set of reason-for-treatment codes was introduced on a trial basis in one of the group practice's clinics. The treating dentist provided a reason-for-treatment code for each of a predetermined set of restorative, endodontic, periodontic, and surgical procedures. The set consists of 11 reasons arranged hierarchically: (1) initial caries, (2) recurrent caries, (3) acute pulpitis/pain, (4) periodontal disease/abscess, (5) third molars, (6) fractured tooth, (7) insufficient tooth integrity, (8) defective/fractured/lost restoration, (9) esthetics, (10) unknown, or (11) other. Due to this scheme's hierarchical nature, third molars extracted for reasons due to caries or periocoronitis would receive a caries or periodontal code, while extraction of asymptomatic third molars would receive a third molar code. The "unknown" code was reserved for reasons where a diagnosis was not possible, and the "other" code was for diagnoses not included on the list. Dentists in this clinic first participated in several informal discussions and trials devoted to developing the hierarchical list of reasons, and then received a summary and detailed lists of the final set of 11 reasons and a list of procedures for which a reason was required. Additional questions and problem solving occurred via e-mail.

Approximately six months after the introduction of the reason-for-treatment protocol in the trial clinic, the reliability of these individual tooth classifications as well as the patient risk classifications were assessed among the nine dentists in this clinic. The practical necessity of maintaining treatment schedules and not unduly inconveniencing patients made a completely crossed design impossible and a balanced pattern of comparisons between all possible pairs of dentists impractical. Hence, eight dentists' classifications were compared to those of a single "nominal standard" examiner who was also a dentist in the clinic. This dentist, whose selection was based on convenience, performed replicate risk classifications for all possible patients receiving examinations and replicate reason-for-treatment classifications for all possible patients receiving treatment requiring reason codes during approximately 40 hours of clinic operation. Thus, final sample sizes and dentist representation were a function of patient availability when the replicate examinations took place.

Replicate risk classifications were performed either before or following new patient or recall examinations. With access to the same diagnostic information available to the original examining dentist, the nominal standard examiner, who had reviewed the classification criteria prior to the initiation of the replicate examinations but had not participated in special training, independently classified a patient's caries risk and periodontal risk. This examiner and any assisting clinical personnel were blind to the prior classifications. Replicate reason-for-treatment classifications were determined under the same conditions, but prior to the initiation of treatment, and separate reasons were indicated for each procedure scheduled for completion during the appointment. The decision to provide treatment was a "given" for the nominal standard examiner, and hence not part of the assessment of agreement in code assignment. Replicate reasons for single

TABLE 1a
Comparison of Dentists' Caries Risk Classifications with Those of a Nominal Standard Examiner

	Nominal Standard Examiner			
Dentists	Low	Moderate	High	Total
Low	30	4	0	34
Moderate	11	17	0	28
High	0	1	3	4
Total	41	22	3	66

Percent agreement=76%; kappa=0.56.

TABLE 1b
Individual Dentists' Agreement with the Nominal Standard Examiner for Caries Risk Classifications

Dentist	A	B	C	D	E	F	G	H	Total
Classifications	1	2	17	8	13	3	9	13	66
Agreements	1	1	14	5	7	3	9	10	50
% agreements	100	50	82	63	53	100	100	77	76

teeth were identified in 30 patients, for two teeth in 20 patients, and for three teeth in one patient. Patients were presented with a verbal description of the study and their verbal consent to participation was sought. The protocol for the study, including the verbal informed consent procedures, had been approved by the institutional review board of the research center affiliated with the group practice.

Reliability was calculated in terms of percent agreement and kappa scores for comparisons of the nominal standard examiner's classifications with those of the other eight dentists combined. Small numbers of comparisons for any one dentist precluded use of kappa scores at this level of analysis.

Results

Tables 1a and 2a compare risk classifications from the eight clinical dentists with those of the nominal standard examiner for 66 patients for caries and periodontal disease risk, respectively. For caries risk classifications, 76 percent of the classifications were the same, yielding a kappa of 0.56. For periodontal risk classifications, agreement was 83 percent, with a kappa of 0.70. In both comparisons, the majority of subjects received low-risk classifications by one or both dentists. Small percentages of subjects received at least one high-risk classification—5 percent for caries, 9 percent for periodontal risk. Among these subjects, identical high-risk classifications were received by three of four subjects for caries risk, and four of six subjects for periodontal risk. Tables 1b and 2b show rates of agreement of individual dentists with the nominal standard examiner. The rates varied from 50 percent to 100 percent for caries, and from 0 percent to 100 percent for periodontal disease, but tended to be similar for individual dentists for the two classifications.

Table 3a shows the cross-classification of reasons for treatment for 73 teeth. Six dentists' classifications are compared with those of the nominal standard examiner. Percent agreement was 74 percent, with a kappa equal to 0.69. Of 19 total disagreements, 12 involved an initial or recurrent caries classification. Table 3b shows the individual dentist rates of agreement, indicating that five of the six clinical dentists disagreed with the nominal standard examiner in their classifications at least once. Percent agreement was 64 percent when two or more teeth were classified in one patient, and 87 percent when one tooth was classified.

Discussion

The extent of agreement between the clinic dentists and the nominal standard examiner reflected in the kappa values is best characterized as "good" for reason for treatment and periodontal risk classifications, and "fair" for caries risk classifications (9). The kappa values are in the same general range found for reliability of WHO ICD-10 mental and behavioral disorder classifications, another discipline where few precise laboratory or other measurements are available to guide classification (10). Our findings also reflect previously reported levels of agreement among dentists for diagnoses and treatment recommendations (7). Of course, the reliabilities reported here do not reflect general agreement among dentists, but rather agreement between a group of dentists and a single dentist who applied the classification criteria in a focused, conscientious manner. The external validity of the results will depend on the typicality of the clinical dentists and the consistency and accuracy of the nominal standard examiner's interpretation of the criteria.

For periodontal and caries risk, reliability was similar for comparisons involving low-risk and high-risk classifications. Also, there were no high/low classification disagreements; all differences involved adjacent categories. It is possible that agreement would have been stronger if the protocol had been more directive concerning how the modifying factors were to be used in assigning patients to risk categories. Anecdotal information suggested that some dentists considered these factors more often than others. Subsequent to this study, the group practice changed the protocol to resolve this problem. For reason-for-treatment classifications, distinctions between initial and recurrent caries accounted for approximately one-fourth

TABLE 2a
Comparison of Dentists' Periodontal Disease Risk Classifications with Those of a Nominal Standard Examiner

	Nominal Standard Examiner			
Dentists	Low	Moderate	High	Total
Low	27	5	0	32
Moderate	4	24	2	30
High	0	0	4	4
Total	31	29	6	66

Percent agreement=83%; kappa=0.70.

TABLE 2b
Individual Dentists' Agreement with the Nominal Standard Examiner for Periodontal Disease Risk Classifications

Dentist	A	B	C	D	E	F	G	H	Total
Classifications	1	2	17	8	13	3	9	13	66
Agreements	1	0	15	8	9	3	9	10	55
% agreements	100	0	88	100	69	100	100	77	83

TABLE 3a
Comparison of Dentists' Reason-for-Treatment Classifications with Those of a Nominal Standard Examiner

	Nominal Standard Dentist											
Dentists	IC	RC	AP	PD	3rd	FT	II	DR	E	U	O	Total
Initial caries (IC)	14	3				1						18
Recurrent caries (RC)	2	10							1		1	14
Acute pulpitis/pain (AP)			2	1								3
Periodontal disease (PD)				0								0
3rd molars (3rd)					0							0
Fractured tooth (FT)		1				8		1				10
Insufficient tooth integrity (II)	1						9					10
Defective, fractured/lost restoration (DR)		1				1		1				3
Esthetics (E)									0			0
Unknown (U)										0		0
Other (O)		1		3		1					10	15
Total	17	16	2	4	0	11	9	2	1	0	11	73

Percent agreement=74%; kappa=0.69.

TABLE 3b
Individual Dentists' Agreement with the Nominal Standard Examiner for Reason-for-Treatment Classifications

Dentist	A	B	C	D	E	F	Total
Classifications	13	23	16	5	11	5	73
Agreements	11	16	12	2	8	5	54
Percent agreement	85	70	75	40	73	100	74

of all disagreements; for another quarter, clinical dentists' lack of attention to the hierarchical order of the classifications may have led to the disagreement. Informal follow-up indicated that three disagreements stemmed from one clinical dentist's misunderstanding of the "periodontal disease" category.

These results suggest that dentists can formulate reasonably reliable risk classifications for caries and periodontal disease, and can specify reason-for-treatment categories from a limited set of choices with the same degree of reliability. However, there is clearly room for improvement in the reliability of these classifications. Thus, the effects of misclassification should be considered in conjunction with all clinical, administrative, and research-related applications of such classification data.

References

1. Bader J, Shugars D, Hayden W, White B. A health plan report card for dentistry. J Am Coll Dent 1996;63:29-38.
2. Imfeld T, Steiner M, Menghini G, Marthaler T. Prediction of future high caries increments for children in a school dental service and in private practice. J Dent Educ 1995;59:941-4.
3. Dodds M, Suddick R. Caries risk assessment for determination of focus and intensity of prevention on a dental school clinic. J Dent Educ 1995;59:945-56.
4. Benn D, Dankel D, Clark D, Lesser R, Bridgewater A. Standardizing data collection and decision making with an expert system. J Dent Educ 1997;61:885-94.
5. Williams R, Paquette D. Periodontal disease diagnosis and treatment: an exciting future. J Dent Educ 1998:62:871-81.
6. Papapanou P. Risk assessments in the diagnosis and treatment of periodontal diseases. J Dent Educ 1998;62:822-39.
7. Bader J, Shugars D. Variation in dentists' clinical decisions. J Public Health Dent 1995;55:181-8.
8. Bader J, Shugars D. A case for diagnoses. J Am Coll Dent 1997;64:44-6.
9. Landis J, Koch G. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-79.
10. Regier D, Kaelber C, Roper M, Rae D, Sartorius N. The ICD-10 clinical field trial for mental and behavioral disorders: results in Canada and the United States. Am J Psych 1994;151:1340-50.